iassist

international association for social science
information service and technology

l'association internationale pour
les services et techniques
d'information en sciences sociales

NEWSLETTER

vol. 1, no. 4

FALL 1977

TABLE OF CONTENTS

1. IASSIST Steering Committee 3
2. IASSIST Action Group Coordinators 4
3. Chairperson's Report 5
4. Treasurer's Report 6
5. First IASSIST West European Meeting, June 27-29, 1977 7
6. "Titles: The Emerging Priority in Bringing Bibliographic Control to Social Science Machine-Readable Data Files [MRDF]"/Sue Dodd 11
7. "An Overview of Display Terminals"/Michael Melton 19
8. "The Issue of Confidential Data: The Need for Formulation of Policy by the Data Archive and Library"/Alice Robbin 20
9. "On Work in Progress: Directory of Directories"/Donald F. Harrison .. 28
10. "Using SPSS MULT RESPONSE to Generate Filtered Marginals"/ Gary M. Grandon 30
11. "A Source Documentation Style Manual"/Richard C. Roistacher 32
12. Data Organization Registry Form (June 1977) 34
13. Book Notices/Kathleen M. Heim 37
14. New Organizations 41
15. Publications 43
16. Upcoming Meetings 44
17. Index 47
18. Membership Application 54

iassist NEWSLETTER vol. 1, no. 4

The IASSIST Newsletter represents an international cooperative effort on the part of individuals managing, operating, or utilizing machine-readable data archives, data libraries and data services. The Newsletter has as its primary intent the dissemination and exchange of information on significant developments in these information centers. Four times a year this Newsletter will report on activities related to the production, acquisition, preservation, processing, distribution, and utilization of machine-readable data carried out by its members and others in the international social science community. Your contributions and suggestions for topics of interest are encouraged and welcomed.

The Newsletter is published four times yearly, as Winter, Spring, Summer and Fall issues. Information for inclusion in the Newsletter should be typewritten and doublespaced. Authorship of all information should be clearly identified (e.g., "submitted by ... [full name, address, and telephone number]"). Action Group reports should include the names of the participants and their national and institutional affiliations. If an individual(s) chaired a meeting, this information should be noted. The format established in the first volume of the Newsletter should be followed when preparing the Action Group report.

Key Title: Newsletter - International Association for Social Science Information Service and Technology

ISSN - United States: 0145-238X

Editor: Alice Robbin, Data and Program Library Service, University of Wisconsin-Madison

Book Notices: Kathleen M. Heim, Library School, University of Wisconsin-Madison

Secretarial: Sue Zimmerman, Data and Program Library Service

All articles for inclusion in the IASSIST Newsletter should be sent to Alice Robbin, Data and Program Library Service, 4452 Social Science Building; University of Wisconsin-Madison, Madison, Wisconsin 53706 USA.

©Copyright 1977 by IASSIST. News and Notes may be freely reproduced; but substantive articles are covered by the copyright.

iassist NEWSLETTER vol. 1, no. 3

IASSIST STEERING COMMITTEE

Carolyn Geda, CHAIRPERSON, Inter-university Consortium for Political and Social Research, P. O. Box 1248, University of Michigan, Ann Arbor, Michigan 48106 USA

Ed Hanis, TREASURER, Social Science Computing Laboratory, University of Western Ontario, London, Ontario, Canada

Manuel Mora y Araujo, SOUTH AND CENTRAL AMERICAN SECRETARIAT, Centro de Investigaciones Sociales, Instituto Torcuato Di Tella, 11 de Septembre 2139 1428 Buenos Aires, Argentina

Sharon Chapple Henry, CANADIAN SECRETARIAT, Data Clearing House for the Social Sciences, 151 Slater, Ottawa, Ontario K1P 5N1 Canada

Naresh Nijhawan, ASIAN SECRETARIAT, Indian Council of Social Science Research, Data Archive, 35, Ferozshah Road, New Delhi-11001, India

Per Nielsen, WEST EUROPEAN SECRETARIAT, Dansk Data Arkiv, H.C. Andersens Boulevard 38 mezz., DK-1553 Køvenhavn V., Denmark

Krzysztof Ostrowski, EAST EUROPEAN SECRETARIAT, Komitet Nauk Politycznych, PAN, Sekcjz Przetwarzania i Archiwizacji Danych, skrytka pocztowa 12, 00-955 Warszawa

Judith Rowe, UNITED STATES SECRETARIAT, Computer Center, Princeton University, 87 Prospect Avenue, Princeton, New Jersey 08540 USA

Joseph Bonmariage, Belgian Archives for the Social Sciences, University of Louvain, SH-2, 1348 Louvain-la-Neuve, Belgium

Evert Brouwer, Amsterdam Social Science Data Archive, University of Amsterdam, Roeterstraat 15, Amsterdam, Netherlands

Ivor Crewe, Social Science Research Council Survey Archive, University of Essex, Wivenhoe Park, P. O. Box 23, Colchester, Essex CO4 3SO, England

Bjorn Henrichsen, Norwegian Social Sciences Data Services, University of Bergen, Christiesgate 15-19, N-5014 Bergen, Norway

Thomas Madron, Academic Computing and Research Services, 243 Grise Hall, Western Kentucky University, Bowling Green, Kentucky 42101 USA

Guido Martinotti, Archivio dati e programmi per le scienze sociali, via G. Cantoni 4, 20144 Milan, Italy

Cees Middendorp, Steinmetzarchief, Kleine-Gartmanplantsoen 10, Amsterdam-C., Netherlands

Ekkehard Mochmann, Zentralarchiv für Empirische Sozialforschung, University of Cologne, Bachemer Strasse 40, 5 Koln 41, Federal Republic of Germany

David Nasatir, Behavioral Sciences Graduate Program, California State College, Dominguez Hills, California 90747 USA

Alice Robbin, Data and Program Library Service, University of Wisconsin, 4452 Social Science Building, Madison, Wisconsin 53706 USA

ACTION GROUP COORDINATORS

CLASSIFICATION

Canada- Mohan Sharma, Humanities & Social Science Library, University of Alberta, Rutherford North, Edmonton, Alberta

Europe- Ekkehard Mochmann, Zentralarchiv für Empirische Sozialforschung, Bachemer Strasse 40, 5 Köln 41, Federal Republic of Germany

United States- Sue Dodd, Data Library, Institute for Research in Social Science, Manning Hall, University of North Carolina, Chapel Hill, North Carolina 27514

DATA ACQUISITION

Canada- Pierre Lacasse, Centre de recherches en aménagement régional, Université de Sherbrooke, Sherbrooke, Québec

Europe- Marcia Taylor, Social Science Research Council Survey Archive, University of Essex, Wivenhoe Park, P. O. Box 23, Colchester, Essex, England C04 3S0

United States- not activated

DATA ARCHIVE DEVELOPMENT

Canada- Laine Ruus, Data Library, Computing Centre, University of British Columbia, 2075 Wesbrook Place, Vancouver, British Columbia V6T 1W5

Europe- not activated

United States- Alice Robbin, Data and Program Library Service, 4452 Social Science Building, University of Wisconsin-Madison, Wisconsin 53706

DATA ARCHIVE REGISTRY

Canada- Lisa Lasko, Canadian Consortium for Social Research, Institute for Behavioral Research, York University, 4700 Keele Street, Downsview, Ontario M3J 1P3

Europe- Joseph Bonmariage, Belgian Archives for the Social Sciences, University of Louvain, SH-2, 1348 Louvain-la-Neuve, Belgium

United States- John Kolp, Regional Social Science Data Archive, University of Iowa, Iowa City, Iowa 52242

DATA DOCUMENTATION

Canada- Dave L. Salley, Management and Central Services Group, Standards Division, Statistics Canada, Tunney's Pasture, Ottawa, Ontario K1A 0T6

Europe- Cees Middendorp, Steinmetzarchief, Kleine-Gartmanplantsoen 10, Amsterdam-C., Netherlands

United States- Sheldon Laube, C.M. Leinwond & Associates, 5010 Stewart Court, College A/C, Maryland 20740

DATA ORGANIZATION AND MANAGEMENT

Canada- Greg Morrison, Social Science Data Archive, Department of Sociology, Carleton University, Ottawa, Ontario K1S 5E6

Europe- Eric Tannenbaum, Social Science Research Council Survey Archive, University of Essex, Wivenhoe Park, P. O. Box 23, Colchester, Essex, England C04 3S0

United States- William Gammell, Social Science Data Center, University of Connecticut, Storrs, Connecticut 06268

PROCESS-PRODUCED DATA

Canada- John DeVries, Social Science Data Archives, Department of Sociology, Carleton University, Ottawa, Ontario K1S 5B6

Europe- Paul Müller, Institute for Applied Social Research, University of Cologne, Greinstrasse 2, 5000-Köln 41, Federal Republic of Germany

United States- Donald Harrison, National Archives (NNR), Washington, D.C. 20408

iassist NEWSLETTER vol. 1, no. 4

CHAIRPERSON'S REPORT

The European IASSIST Action Group Workshop was held at the Danish Data Archive in Copenhagen, June 26-29, 1977, under the direction of IASSIST West European Secretary, Per Nielsen. A major product from this meeting was a data organization registration form (DORF) which is reproduced in this Newsletter. Copies of the DORF can be obtained by writing to Elliott Avedon, Department of Recreation, University of Waterloo, Waterloo, Canada. The completed form should be returned to him. A study description questionnaire form developed by the Zentralarchiv at Cologne was reviewed and discussed. Several archives are now testing this instrument. Individuals interested in examining or testing the instrument should contact Elliott Avedon for copies. The Workshop also addressed levels of documentation for machine-readable files. The results of this discussion are reported in this Newsletter.

The program for the February 8-11, 1978, IASSIST Conference at Nordic Hills, Itasca, Illinois has been revised on the basis of membership feedback. The following panels have been scheduled: Documentation, Alternative Structures for Data Access, Problems and Potentials of Networking, Privacy versus Freedom of Information, Acquisition and Preservation, and Software Analysis of Non-Rectangular Files. In addition, workshops will be offered and action group meetings will be held. Although panelists have been designated for each session, the sessions have not been closed. Individuals interested in presenting a paper at any of the sessions should send the proposed title with a 100 word abstract to Carolyn Geda. Please note that individuals who were following the original program theme and who anticipated doing a formal presentation within an Action Group should submit their paper to the Chairperson. The schedule for the panels, workshops and action group meetings will be sent to IASSIST members with conference registration forms during November. Please note that the conference will terminate on Saturday, February 11, 1978. Additional information on the February meetings is contained in the IASSIST Newsletter section on "Upcoming Meetings."

The panels for the IASSIST sessions at the International Sociological Association World Congress are Privacy versus Freedom of Information (chairperson Guido Martinotti), Issues in Comparative Data and Research (chairperson Elliott Avedon), and Research Problems Associated with Complex Data Bases (chairperson John De Vries). Individuals interested in presenting a paper should contact the chairperson as soon as possible. Information on the availability of charter flights will be sent to IASSIST members in the near future.

This issue of the Newsletter completes the first volume and also contains an index for the four issues in this volume. The first volume has been edited and produced by Alice Robbin, University of Wisconsin-Madison. The effort she has devoted to the Newsletter has been widely recognized and appreciated. She will remain affiliated with the Newsletter in an ex officio capacity. Articles and notices to be included in the Newsletter should continue to be sent to her. The IASSIST Newsletter will move to Western Kentucky University where Thomas Madron, Academic Computing and Research Services, will assume the role of editor.

A nominating committee will be appointed shortly to prepare a slate for the election of the Steering Committee members. If you are interested in serving on this committee or on the Steering Committee or wish to recommend other individuals or nominations, please forward the names to Carolyn Geda.

Ed Hanis, treasurer, has written a treasurer's report for the period January-July 1977, which is reproduced in this Newsletter. IASSIST is financially solvent at the end of its first formal year. The two conferences sponsored by the Canadian and U.S. Secretariats contributed greatly to the financial success of the first year. There is some justification for increasing the membership fee; however, the decision to maintain the fees at the current level has been made on the basis of continued membership growth and decreased printing costs for the Newsletter, as a result of moving the Newsletter to Western Kentucky University. Although the Secretaries have been successful in soliciting memberships, we ask that you join us in a membership dirve. A publicity flyer is being produced in multiple copies and will be available through the Secretaries.

TREASURER'S REPORT

This report covers all receipts and expenditures since the commencement of IASSIST's financial life through July 31, 1977. The Cash on Hand as of July 31 should be sufficient to cover remaining expenses for calendar year 1977, including the cost of printing and distributing Nos. 3 and 4 (Volume I) of the Newsletter.

	Amount in Can. or U.S. Dollars
A. Receipts	
1. Membership Fees (for 1977)	$2,665.00
2. Workshop Registration Fees (Cocoa Beach and Toronto)	1,640.00
3. Interest, bank accounts	5.64
Total Receipts	$4,310.64
B. Expenditures	
1. Newsletter (Vol. 1, Nos. 1 and 2)	$1,114.93
2. Members' Subscriptions to SS Data, Vol. VII, Sept. 77 to June 78	411.00
3. Workshop Expenses	535.13
Total Expenditures	$2,061.06
C. Cash on Hand (as of July 31, 1977)	$2,249.58

iassist NEWSLETTER vol. 1, no. 4

FIRST IASSIST WEST EUROPEAN MEETING
june 27 – 29, 1977

Report of the West European IASSIST Workshop Held in Copenhagen, June 27-29, 1977

Submitted by

Per Nielsen, Danish Data Archives

Members Present

Europe
- Tomasz Bankowski, Zowar Computing Centre, Warsaw, Poland
- Flemming Bigom, Danish Data Archives, Copenhagen, Denmark
- Merete Watt Boolsen, Danish Data Archives, Copenhagen, Denmark
- Ulf Christoffersson, University of Gothenburg, Sweden
- J.C. Deheneffe, University of Louvain-la-Neuve, Belgium
- Bartlomiej Gasiorowski, Polish Academy of Sciences, Warsaw, Poland
- Jørgen Grosbøl, Danish Data Archives, Copenhagen, Denmark
- Astrid Bogh Lauritzen, Danish Data Archives, Copenhagen, Denmark
- Cees Middendorp, Steinmetzarchief, Amsterdam, the Netherlands
- Ekkehard Mochmann, University of Cologne, Federal Republic of Germany
- Per Nielsen, Danish Data Archives, Copenhagen, Denmark
- Krzysztof Ostrowski, Polish Academy of Sciences, Warsaw, Poland
- Karsten Boye Rasmussen, Danish Data Archives, Copenhagen, Denmark

North America
- E. M. Avedon, University of Waterloo, Ontario, Canada
- Carolyn Geda, University of Michigan, Ann Arbor
- Sharon Chapple Henry, Data Clearing House for the Social Sciences, Ottawa, Canada

Agenda

June 27: What do we need to know about each European Archive? How much detailed information is really required? Can we design one questionnaire that will include all of the required elements and not be a burden to the respondent?

June 28: Should archive holdings be standardized with respect to documentation? If so, what should be included in these standards? How can we ensure that these standards are adhered to and used by primary research personnel prior to primary analysis?

June 29: What are the sources of data? Should Archives classify and index these data in a standard manner to facilitate data retrieval as well as cross-national and international research? If so, how should this be accomplished?

On behalf of the West European IASSIST Secretariat and the DDA, Per Nielsen welcomed the participants and expressed his gratitude that so many "guest invitees" from the East European and North American Regions were present. To bring everybody up-to-date, information was provided on the latest "data-conferences": Carolyn Geda gave information about the IASSIST North American Conference in Florida (February 1977), where 35 data information professionals had attended. Carolyn Geda and Sharon C. Henry then discussed the second North American IASSIST meeting in Toronto (May 1977), in which the number of attendees had doubled compared to the Florida meeting. Working papers and ideas from the above meetings were useful inspiration throughout the Copenhagen meeting. Ekkehard Mochmann reported on the Moscow Conference on Information and Documentation in Social Sciences from which he had just returned; the meeting was an initiative of the Vienna Centre, hosted by INION and supported by UNESCO. The upcoming IASSIST meetings in Chicago (February 1978) and Uppsala (August 1978) were described.

The issue of the first day in Copenhagen was discussed at length. There was agreement that it would be possible to design a Data Organization Registry Form that included all the elements required for a full description of data service organizations. Consequently, it was decided to construct such a Data Organization Registry Form and to recommend this form to IFDO for use; the intended audience for the form was to be the administrators of the various data organizations.

As source material for the Data Organization Registry Form the following documents were used:

1. Questionnaire on data acquisition policies and problems (Marcia Taylor, SSRC Survey Archive, Essex)
2. Questionnaire on data preparation procedures (Eric Tanenbaum, SSRC Survey Archive, Essex)
3. Preliminary list of data elements and subject terms for IASSIST Data Archive Registry in Canada (Lisa Lasko, Institute for Behavioral Research, York University; and Lana Prokop, University of Toronto)
4. Preliminary outline for a "A Guide to Providing Social Science Data Services" (Alice Robbin, University of Wisconsin; and Laine Ruus, University of British Columbia)
5. Request from Alice Robbin and Laine Ruus regarding service documents
6. Input Form (Directory of data bases in the social and behavioral sciences)
7. Data Documentation Form (Data Clearing House for the Social Sciences)

The Group constructed a preliminary questionnaire form consisting of material selected from the above documents; then modified, added, and deleted; and finally ended up with a handwritten version to be computerized before a final review during the Wednesday session.

The working group made an attempt to set up a list of elements describing a data file that would be useful from the user point of view:

(i) Library information (such as card catalogue)
(ii) Archive information/Study abstract
(iii) Study description (in machine-readable form)
(iv) List of variables (in machine-readable form)
(v) Codebook (in machine-readable form)
(vi) Classification/index (in machine-readable form)
(vii) Data (not part of documentation, but the basis for the file)
(viii) Special publications (either in print or machine-readable)

Each of the documentation items were subjected to detailed discussions, and some recommendations were agreed upon; a summary is listed below:

(i): Library Information: The North American IASSIST Classification Action Group (coordinated by Sue A. Dodd, University of North Carolina) has set standards for cataloging machine-readable data files. Data organizations and primary research personnel should adopt the library information recommendations of this Action Group.

(ii): Archive information/Study abstract: Each data organization has its own way of publishing its data holdings in an inventory; it seems difficult to make recommendations regarding production of study abstracts. However, it is recommended that data organizations as well as research institutions produce information at the abstract level.

(iii): Study description: It is recommended that the standard study description scheme (developed as a result of a meeting in Copenhagen in June 1974) be tested by several data organizations. As a concrete step in that direction several data holders agreed to test this instrument during the next year:

BASS, Louvain-la-Neuve, Belgium
DDA, Copenhagen, Denmark
ICPSR, Ann Arbor, Michigan, USA
LSDB, Waterloo, Ontario, Canada
Steinmetzarchief, Amsterdam, the Netherlands
Zentralarchiv, Cologne, Federal Republic of Germany
Ulf Christoffersson, University of Gothenburg, Sweden

Five study description questionnaire forms will be disseminated to each of the above holders along with an instruction in the use of the standard study description scheme. Each data holder will send study descriptions (in English language) of five different files to the Zentralarchiv. Preferably, the study descriptions should be submitted on tape; however, the ZA would accept filled-in questionnaire forms from data holders unable to make the study descriptions machine-readable. Attempts will be made to furnish the necessary software for printing the study description in the participating institutions. The outcome of this testing will form the basis for an extension and updating of the instruction manual which is presently being developed and distributed from the DDA. At the end of this process, final recommendations can be made.

(iv): List of variables: It is recommended that the list of variables be available in print as well as in machine-readable form.

(v): Codebook: The "ideal" codebook was outlined by the workshop as consisting of the following 15 items (the first three items referring to file level, the last twelve to be supplied variable by variable where applicable):

(1) Title of study (file and subfile names)
(2) Format of the data-file
(3) Comments at file level (e.g. concerning application of missing data codes; special weighting features; special precautions for use)
(4) Variable identification (number; label; short name; mnemonics)
(5) Variable source reference
(6) Variable location and length
(7) Variable type (alpha; alphameric; numeric; symbolic)
(8) Number of decimal places (scale of measurement)
(9) Source statements/texts/questions/scale descriptions/introductory statements related to the responses

(10) Code values
(11) Code descriptions
(12) Comments: coding and field work (coding and interviewer instructions)
(13) Variable contingencies (filter; skip; probe; control questions)
(14) Variable consistency
(15) Derived variables

(vi): Classification/index: It is recommended that future IASSIST meetings be concerned with this topic. A number of classification schemes (on file and variable level, respectively) are available; however, further testing and elaboration is required before final recommendations can be made.

(vii): Data: Single punch data is preferable for archiving purposes.

(viii): Special publications: It was recognized that many data archives produce publications concerning specific files (examples were examined in hard copy). There was some discussion regarding machine-readable special publications and the future directions that this area of documentation may follow.

The machine-readable preliminary Data Organization Registry Form was reviewed. Questions were added, deleted, modified. This work was finished with agreement on a Data Organization Registry Form to be disseminated to the participants for completion; comments arising from the application of the Form will be reported to the DDA; such comments may make yet another editing process inevitable.

The working group discussed sources of data. This topic had been subject to discussion during the Data Organization Registry Form and Data Documentation sessions of the preceding days; however, the classification/indexing problems regarding data from various sources could not at this meeting be operationalized to a level allowing for recommendations. On the contrary, it is evident that there is some confusion right down to the level of terminology. It would be very useful if the Classification and Process-Produced Data Action Groups of IASSIST could come up with a tentative taxonomy, taking into consideration the increasing number of different data sources relevant to the social science community.

Summary and Prospect

(1) The Data Organization Registry Form will be disseminated from the DDA to the participants as soon as final editing is completed.
(2) No later than October 1st participants will complete the Form on behalf of their organizations and return the completed version to the DDA with any comments and suggestions that may arise from the respondent role.
(3) The Data Organization Registry Form will be recommended to the IFDO for use.
(4) Concerning the Study Description Scheme participants will receive a manual (Applications & Instructions).
(5) Participants who have agreed to take part in the testing of the study description will receive five Study Description Forms to be completed in English for five different files in their holdings.
(6) No later than December 1st, the completed study descriptions will be transferred to the ZA (preferably in machine-readable form) along with a short report on the test experience.

iassist NEWSLETTER vol. 1, no. 4

DISCUSSION PAPER / sue dodd

Titles: The Emerging Priority in Bringing Bibliographic Control to Social Science Machine-Readable Data Files [MRDF]

by

Sue Dodd
Institute for Research in Social Science
University of North Carolina
Chapel Hill, NC

Two recent attempts to compile "catalogs" of social science data have encountered the lack of consistency among titles for the same data set. One attempt has been the recent cataloging efforts at the Universities of North Carolina, Wisconsin, Princeton, and Yale, whereby traditional library cataloging records are created for social science data generated by academic research. The other has been the efforts of the Association of Public Data Users (APDU) to compile a directory of publicly available data files which represent primarily government produced data. Both groups have experienced the same problem: variance of titles for the same data file. Yet, without some control over titles and some mutually agreed upon primary source of title information, there can be no bibliographic control of social science data and none of the related products such as a union list of machine-readable data files. This paper will attempt to offer some suggestions for remedying the situation, including guidelines for transcribing titles; for creating a "title page"; for compiling a bibliographic reference; and for establishing an "authority list" for titles.

Origins of titles for social science data files

Unlike a book a social science data file may exist for a long time without a title. Until it has been properly titled, it may be known only by a study number (*e.g.*, Study #5063), or by the name of the principal investigator (*e.g.*, The Stouffer Study), or by the source of production (*e.g.*, The RAND Survey). If a data file survives the time period between data collection, data analysis, and data publication sans title, it is likely to assume the title of the primary publication (*e.g.*, *Communism, Conformity, and Civil Liberties*).

The first appearance of a title usually occurs with the generation of early sources of documentation. Documentation may include a questionnaire, coding instructions, codebook, manual, or project report. Given the nature of MRDF, some type of accompanying documentation is required in order to "read" the data. Titles recorded on documentation are also the most visible because "containers" (protective canisters) of MRDF have no identifying titles; or if they do, it is usually a shortened title given the space constraints of the container. However, an initially applied title of a MRDF is not necessarily the only or lasting one.

During the life cycle of a social science data file, a title is frequently changed or modified as responsibility changes for file creation, processing, analysis and reporting. For example, one group of persons may be responsible for actual data collection plus the conversion to an "automated" format, while another group may be responsible for the data analysis and data reporting. Such diversification of labor often leads to different titles for the same data. After the primary analysis, reporting, and possible publication by the principal parties, a data file may be deposited with a data archive, center, or library for the purpose of secondary analysis. At this point, the data and documentation may go through further processing, including a new codebook and a new title. At about the same time, but not necessarily by the same person, a data abstract, study description or some type of informational notice is written to publicize its availability to the general public.

If there are dual or multiple distributors of the same data file, titles could easily vary from distributor to distributor. For example, there are at least three known distributors for a particular Harris survey with the following titles:

Violence in America
The American Public Looks at Violence
Harris 1968 Violence Survey, #1887
Harris Poll: "The American Public Looks at Violence"

In the case of the APDU directory, the overlap of mutually held and accessable public data files was impossible to determine, since members had listed the same data file under various titles. For example:

City and County Data Book
County and City Data Book
1972 County and City Data Book
County and City Data Book Tape

The cataloging experience at the University of North Carolina has revealed that out of approximately 500 separate social science data files from many different sources, close to 80 of these files had variant titles. In most cases, the title in the codebook varied from informational listings provided by the distributor of these data.

As the life cycle of a data file continues, popularized titles begin to evolve and grow organically and are usually a modified version of the primary title. For example:

Modified title: French and German Elite - Arms Control Data
Primary title: Arms Control in the European Political Environment: French and German Elite Responses, 1964

Other titles are compressed into acronyms:

Modified title: The CSEP Study
Primary title: The Comparative State Election Project

Others take on the name of the principal investigator(s):

Modified title: The Matthews-Prothro Study
Primary title: The Negro Political Participation Study

Finally, variant titles may appear in "notes" or in bibliographic references in the various scholarly journals. Without any guidelines on how to cite numerical data files and without any control over the proliferation of titles, title information will vary among scholars. Often, the fault rests not so much with the person citing the data as it does with the distributing agency which has failed to provide proper bibliographic information on a particular data file.

Summarizing, the history of a data file usually reveals the various levels of title changes and modification. However, it is unlikely that a cataloger or a scholar will have access to this history. Instead, he will be confronted with the problem of choosing or citing one title from among many for his respective uses.

Guidelines for transcribing titles

In our attempt to offer suggestions on how this situation could be remedied, this section describes the basic components of a title and suggests guidelines on how to transcribe a title for social science data.

Components of a title for social science data files would include the following: 1) descriptive words indicating content; 2) geographic focus or unit; 3) chronological year(s) of data target or data collection; 4) source of data (e.g., court records); 5) producer, contributor or sponsor of data; and 6) study or series number (if important for ordering or for distinguishing one data file from another).

Guidelines include the following:

I. Make the title as descriptive and as complete as possible. A good title should be descriptive of the contents of the document or data it is describing and should include as many of the components described above as are applicable. If there is one major theme or focus, then this should be mentioned in the title. If the data contain information on many different topics, none of which appears to dominate, then a broader or more general subject approach may be taken (e.g., Harris 1972 Public Opinion Survey; or the National Opinion Research Center 1974 General Social Survey; or the Survey Research Center 1976 Social Indicator Survey). When transcribing a title, be aware that the descriptive words contained in a title take on added significance with the existing technology for keyword or full-text retrieval. For example, the only subject approach to SOCIAL SCIENCE CITATION INDEX (whether it be by the printed reference work or by the on-line search capability) is via the descriptive words contained in the respective titles.

II. If at all possible, DO NOT take a title from a publication based on the data file, as this may cause copyright violations and problems with international coding schemes such as ISBN (International Standard Book Number). If this cannot be avoided, then a qualifier should be attached at the end of the title. For example:

Civic Culture (Machine-readable data file)
Communism, Conformity and Civil Liberties (MRDF Source Documentation)

III. For any data that are part of a predictable series (occurring at definite time intervals, such as the Census or election surveys), titles should be consistent throughout the life of the series.

IV. For data that are part of an on-going collection or series (collected at non-predictable intervals and with varying subject focus), one may consider the following sequential title arrangement: 1) organizational name of producer; 2) chronological date of data target or data collection; 3) geographic focus (if unique); 4) descriptive content (including subtitles); and 5) study or series number.

An on-going series of data (such as public opinion polls) tends to be associated with the originating source or producer of these data. Therefore, it is recommended that the organizational name of the producer come first. This arrangement also allows for a large collection of data from the same source to be grouped alphabetically for easy reference. For example:

Harris 1969 Morals and Values Survey, No. 1933
Harris 1969 Science, Sex and Morality Survey, No. 1927

In cases where there is more than one data collection per year on a given topic, the month or season could follow the year in parenthesis. For example:

Survey Research Center 1957 (Fall) Consumer Attitudes and Behavior Survey, No. 3631

If the geographic focus is unique, it is recommended that it be included in the title. For example:

American Institute of Public Opinion 1975 Japanese Election Survey, No. 7811
Harris 1965 Dallas Sports Survey, No. 1545

To indicate that data in a continuing series may have a varying subject focus, it is recommended that sub-titles be used. For example:

Survey Research Center 1963 Detroit Area Study: A Study of Family-School Relationships
Survey Research Center 1964 Detroit Area Study: The Measurement and Validation of International Attitudes

Study numbers should be included in the title if they are part of an on-going collection of data and are consequently helpful in distinguishing one data file from another. For example:

Harris 1967 Public Opinion Survey, No. 1702
Harris 1967 Public Opinion Survey, No. 1718

National Opinion Research Center 1963 (January) Amalgam Survey, SRS-100
National Opinion Research Center 1973 (December) Amalgam Survey, SRS-4179

V. Avoid beginning a title with articles (such as a, an, the, etc.).

VI. Avoid beginning a title with numerics (e.g., 1972 County and City Data Book). With most computerized alphabetic listings, those titles beginning with numerics are placed either at the very beginning of a listing or at the very end. Such placement may cause certain data files to be overlooked.

VII. Avoid using acronyms in titles. The full meanings of acronyms should be spelled out and if used at all, should follow full meanings enclosed in parenthesis. For example:

World Event/Interaction Survey (WEIS)

VIII. When applying titles to sub-sets of data files, indicate both the original data title and the fact that it is derived from a larger file. For example:

Comparative State Election Project: Federal District Sub-File

Title control and an "authority list" of titles

Title control for social science data must be applied at one of two stages in the life of a data file: either at the production level or at the distribution level. Ideally, the creator or producer of a MRDF should apply the "authoritative" title. However, in those cases where this responsibility has (for whatever reasons) defaulted to the distributor of the data, then he should provide the singular title. All other references to this MRDF should carry this title.

One way to bring some order to the existing chaos among titles, is to establish an "authority list" of titles for social science data files. Such an effort is being undertaken by members of APDU; and a similar "union list" of catalog records would have the same effect. Again, the primary responsibility for establishing or determining the authoritative title rests first with the producer and then with the distributor. If the producer has abdicated that responsibility when depositing a data file with an archive or data center, then responsibility lies with the distributor.

In those cases where there are multiple distributors of the same data, then a determination has to be made as to the one with the most "authority," or official status, or national prestige, etc.

For data files that have been changed through major processing techniques or reformatted for a more efficient "reading," or have been changed in terms of content or observations, then the title remains the same but the data become a new edition. Thus, an "authority list" of titles would include the various editions of MRDF, just as the National Union Catalog (NUC) carries the various editions of books.

Sub-files taken from larger data files should carry a distinctive title, and if not, the producer or distributor should modify the title with some type of qualifier (e.g., sub-file A; selected sub-files, etc.).

The major data producers and distributors of social science data would be responsible for publishing their respective lists of authoritative titles. These "authority lists" of titles could then be published in some appropriate newsletter or publication such as SSdata or the IASSIST Newsletter.

In establishing "authority lists" of titles there is also the need for establishing a concensus on the primary source of title information. If there is a title on the codebook and another title in a directory, which is the correct title? Without having access to the history of a particular data file, how can the judgement be made as to the proper title? One answer would be to rate, in order of importance, the various sources of documentation. For example, the codebook or its equivalent would be the primary source of title information; the data directory or study description would be the secondary source; the reporting source or publication would be the third, etc. Some discipline has to be applied to social science documentation in general, but specifically to the bibliographic aspects of such documentation. Such responsiblity should not end with titles, obviously, but should be extended to include all the components of a bibliographic citation. For example, the information required for compiling a bibliographic reference should come from the documentation accompanying a data file, and the most obvious place to derive this information would be from the "title page" of that documentation.

Title page for social science MRDF

In the past, very few data producers or distributors have taken the style or content of the title page of documentation seriously. However, if social science data files are to be readily accepted into the mainstream of bibliographic control, then more attention has to be given to these title pages. For example, the title page of a book becomes for the cataloger, the principal source of information. It is so respected by catalogers that the information contained on the title page becomes as "dogma" and cannot be deviated from in the transfer of information to the catalog record. However, the quality and amount of information provided on most title pages of social science documentation cannot be taken seriously by a cataloger. In truth, many sources of documentation for MRDF have no title page equivalent.

Information contained on a title page of MRDF documentation should consist of the basic bibliographic components including authorship; title; medium designator; edition; imprint; and series statement.

The "medium designator" is a term used to denote the generic form or type of material listed or referenced. It is used to distinguish one type of medium from another and to provide clarity. The most universally accepted term for this medium is "machine-readable data file".

The "imprint" includes the place of production; name of the producer; date of production; place of distribution; and name of distributor. The producer is defined as that party responsible for the collection, compilation, and physical production of the data (*i.e.*, the mechanized process of transforming information into the format known as "machine-readable") and the distributor as that party responsible for disseminating the data to others upon request.

The "series statement" would provide the reader with relevant information about an on-going collection of data (*e.g.*, SRC/CPS 1958 American National Election Study, No. 4).

As mentioned earlier, an "editon" occurs when data files have been modified through major processing techniques or reformatted for a more efficient "reading" or when the data have been changed in terms of content or observations. Edition statements appear in an abbreviated format on a title page (*e.g.*, DUALabs ed., NORC rev. ed., or 1st ICPSR ed.).

Although, it is highly recommended that only the basic bibliographic information be placed on a title page, there may be situations where additional information would be helpful. Examples would include: date of data focus, if not part of the title; source of funding, if appropriate; scope of documentation, if documentation consists of more than one volume; study number, if necessary for identification or ordering purposes; etc.

Information pertaining to unique classification schemes such as the International Standard Book Number (ISBN); the Library of Congress Card Number; and the catalog card facsimile should appear on the verso of the title page. For an example of a title page of a machine-readable data file codebook, see appendix.

Placement of the information is flexible. However, placement of a study number behind or immediately under a title will be construed as being part of that title. For example:

The SRC 1952 Election Study (S400)
or
The 1972 German Election Panel Study
(Zentralarchiv Nos. 635,636,637 -- ICPR No. 7102)

Bibliographic references for social science data files

The title page, in addition to providing the basic information required for the catalog record, should also provide the information required to compile a proper bibliographic citation. The Classification Action Group of IASSIST has been working to define the necessary components and structure for a proper bibliographic citation and has followed the guidelines provided in the forthcoming publication entitled: American National Standard for Bibliographic References. However, these guidelines have been modified slightly to represent the particular needs of social science numerical data and to conform with the forthcoming AACR II cataloging rules.

The basic components of the bibliographic reference would include authorship; title; medium designator; sub-title; edition; imprint (place of production, name of producer, date of production; place of distributor, and name of distributor); extent of file; notes; and series statement. The examples that follow were provided by the Classification Action Group:

Title first: Mexico's naturalized citizens, 1828-1931 [Machine-readable data file]. Principal investigators: Harold Sims, Susan Sanderson, and Philip Sidel. Pittsburgh, PA : University of Pittsburgh, 1975-76 [producer and distributor]. 1 data file (8,066 logical records).

Author first: Shanas, Ethel. The health of older people [Machine-readable data file] : a social survey : public attitudes of older people. Norc rev. ed. Chicago : National Opinion Research Center, 1957 [producer and distributor]. 1 data file (2567 logical records) and accompanying codebook (166p.).

Swidzinski, Susan. Syllabication [Machine-readable data file] : a drill and practice lesson. Bloomington, MN : Control Data Corporation, 1976. On-line program lesson available only via the Plato System.

Henry, Neil. MAXCLS.BAS [Machine-readable data file] : a program for maxium likelihood estimation of paramaters of unrestricted latent class models. Lafayette, IN : Gary Income Maintenance Experiment, 1974 : Pittsburgh, PA : Social Science Computer Research Institute [distributor]. 1 program file (95 statements, BASIC) and accompanying manual (53p.).

The information contained in the brackets, while highly recommended by the Classification Action Group, are optional according to the ANSI standards and the forthcoming AACR II rules. The extent of file (logical records, program statements, etc.) and "notes" are also optional. Many distributors of data already provide the user with some "data acknowledgement" information. Guidelines or examples of how to cite these data in the literature should be part of this information.

Conclusion

In conclusion, we have discovered that it is not unusual for social science data files to receive many different titles in their lifetime. Some titles are modified through data processing efforts; others grow or evolve from popular usage; and even others are erroneously recorded from one source to another. This lack of title control has proven to be detrimental to the efforts of those who are attempting to compile any authoritative listing or "catalog" of social science data. It is also apparent that there is an immediate need for an "authority list" of titles, and that some decision has to be made regarding the primary source of title information.

It is hoped that this paper will bring these problems and needs to the attention of those parties who have both the responsibility for and the control over the situation. Critical attention and immediate action, on the part of the major producers and distributors of data, is necessary to bring about true title control and better bibliographic documentation. Without it, information on social science data files will remain in "elite obscurity".

Appendix

MACHINE-READABLE DATA FILE CODEBOOK

BERKELEY RADICALS FIVE YEARS LATER: A FOLLOW-UP
SURVEY OF STUDENTS WHO WERE ARRESTED IN THE 1964
FREE SPEECH MOVEMENT AT THE UNIVERSITY OF CALIFORNIA

Conducted by
The Detroit Free Press of Knight-Ridder Newspapers, Inc.
under the direction
of
Philip Meyer and Michael Maidenberg

July 1969

SSDL 1972 rev. ed.

SOCIAL SCIENCE DATA LIBRARY - UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA
27514

iassist NEWSLETTER vol. 1, no. 4

DISCUSSION PAPER/ michel melton

An Overview of Display Terminals

Michael Melton
Battelle Columbus Laboratories
Columbus, Ohio

Overview

Display terminals can be generally classified into three categories: (1) dumb, (2) semi-stupid, and (3) intelligent.

Dumb terminals offer a limited number of functions and most closely resemble Teletypes. Semi-stupid terminals usually offer a certain amount of features, most likely, data input editing and formatting. Some terminals in this class can be tailored to fit a specific application via a limited programming capability. Intelligent terminals are supported by software programs. The vendors typically provide an operating system, an assembler or a compiler, I/O utilities, and one or more application programs such as text editing or data entry. Programmable terminals provide the user with the highest degree of flexibility, permitting the terminal to be tailored to the user's environment. Some programmable terminals are available as turnkey terminals, which means that the program is furnished by the vendor and the terminal is ready to use as soon as it is installed.

Cost is usually proportional to capability. Dumb display terminals are the least expensive at $1000 to $2000. Intelligent terminals range upward from $6000. Expensive items include increased memory, additional display units, and peripherals such as diskette or disk storage and printers. Semi-stupid terminals are priced somewhere inbetween.

Some brand names in dumb terminals include Applied Digital Data Systems, Beehive, Infoton, and Lear Siegler. Some of the semi-stupid terminals include Hewlett-Packard 2640, IBM 3270, ICC 40+, and the Univac Uniscope. Some intelligent terminals are Applied Digital Data Systems 70, the Beehive B800, Datapoint, Four-Phase Systems Incoterm, Sycor, and Univac UTS400.

Most terminals introduced on the market in the last three years have been microprocessed-controlled (i.e., intelligent). Microprocessors are cheap, cut design, development, and production costs, and they lend themselves to a variety of applications that can be implemented by the vendor or the user. Computer programs which control the functions of an intelligent terminal are called firmware. The user can control the terminal's functions by either changing a set of parameters or by adding new firmware to the microprocessor.

Display Functions

Some different terminal display functions are:

(1) Color - Few display terminals offer color but some offer up to eight colors
(2) Reverse video - A negative image of the data, data normally displayed in white or dark background is displayed in black or white background
(3) Programmable brightness level
(4) Character or field blinking

(5) Roll or scroll data is rolled up or down the screen, permits the user to scan a large volume of data
(6) Paging - Data is stored on pages (a full screen) user is able to review any selected page.

Editing Functions

Editing features include:

(1) Character deletion
(2) Line insertion
(3) Line deletion
(4) Erase
(5) Character repeat.

External I/O devices can add flexibility to the applications possibilities for display terminals. A cassette tape drive or diskette drive can be used to store display formats, data to be transmitted, or user programs. A printer can provide hard copy.

Selecting a Terminal

Some questions you should ask yourself when selecting a display terminal are:

(1) What are the essential parameters for a display terminal that will satisfy your needs?
(2) Who supplies the terminals with the features you desire?
(3) Maintenance provisions?
(4) Talk to users concerning problems encountered when installing it, failures that have occurred, and any incompatibilities.

DISCUSSION PAPER / alice robbin

The Issue of Confidential Data: The Need for Formulation of Policy by the Data Archive and Library

by

Alice Robbin
Data and Program Library Service
University of Wisconsin-Madison

The Issue of Confidential Data

I.

During the last decade there has been increased concern about the problems of confidentiality involved in the collection and dissemination of individual microdata. Concern has revolved around the government's perceived need to collect increasing amounts of information at a microdata level for social policy formulation and evaluation, types of information which potentially compromise personal privacy; lack of good security measures for protecting access to personal information; availability of this information to individuals, private firms and public agencies unrelated in any way to the primary producers of the statistical data; growing restrictions placed on the access to statistical microdata by the statistical agencies, restrictions which appear to be in response to public pressure; and, the impact of restrictions on access to these data which are needed for research and public policy development by the scholarly community and others engaged in statistical analysis. This concern has been manifested in a very large literature, formal conferences and commissions, and legislation.[1] Recently, a meeting concerning the issues of privacy, confidentiality, and the use of government microdata for research and statistical purposes was held at the Rockefeller Foundation's Bellagio Study and Conference Center at Lake Como, Italy, August 16-20, 1977[2].[3] It seems useful to summarize the draft Report on this conference because of the importance of this issue for the data archive.

II.

David Flaherty, author of the draft Report describes the origins of the Bellagio meeting. During the last three years he and Edward Hanis of the University of Western Ontario have been involved in the Privacy Project studying the problems of privacy and confidentiality involved in the collection and dissemination of individual microdata by central statistical agencies in Canada, the United States, the Federal Republic of Germany, Sweden, and the United Kingdom. The Privacy Project identified key personnel in each central statistical agency and found that they face common problems in the formulation of policy on confidentiality and data dissemination. The Privacy Project found that although these individuals face common problems they have had surprisingly little contact with one another on these issues. The Bellagio Conference brought together leading individuals from each of the five statistical agencies to discuss common problems and solutions. The Conference included sociologists, economists, and historians because it was felt necessary to demonstrate to the "custodians" of government data that a real need exists for access to individual microdata for research and statistical purposes. A series of general principles or propositions pertaining to research and statistical uses of individual information held by government agencies evolved during the meeting.

I have reproduced the 18 principles in toto, since I feel it useful to add to the public debate on the privacy/confidentiality issue. Many of the principles, while of potential interest to the data archive community do not directly affect data archivists at a local repository level; however, I have starred (*) those which I think may prove applicable to the archive's role.

1. *National statistical offices should provide researchers both inside and outside government with the broadest possible access to information within the bounds of accepted notions of privacy and legal requirements to preserve confidentiality.*

2. *Legal and social restraints on the dissemination of microdata are appropriate when they reflect the interests of the general public in an equitable manner. These constraints should be re-examined when they result in the protection of vested interests or the failure to disseminate anonymized information for statistical and research purposes without direct consequences for a specific individual.*

3. *All copies of government statistical records should be rendered immune from compulsory legal process by statute.*

4. *In making data available to researchers national statistical offices should provide some means to ensure that decisions on selective access are subject to independent review and appeals.*

*5. *The distinction between a research file in the sense of a statistical record (as defined in the 1977 report of the U.S. Privacy Protection Study Commission), and other micro files is fundamental in discussions of privacy and dissemination of microdata. All dissemination of government microdata discussed in connection with the Bellagio Principles is assumed to be a transfer of data to research files.*

6. *There are valid and socially-significant fields of research for which access to microdata is indispensible. One of the prime sources is government microdata, including statistical agencies.*

*7. *Public use samples of anonymized individual data are one of the most useful ways of disseminating microdata for research and statistical purposes.*

*8. *Techniques now exist that permit preparation of public use samples of value for research purposes within the constraints imposed by the need for confidentiality. Countries with strict statutes on confidentiality have prepared public use samples.*

*9. *There are legitimate research purposes requiring the use of individual data for which public use samples are inadequate.*

*10. *There are legitimate research uses which require the utilization of identifiable data within the framework of concern for confidentiality.*

*11. *Other techniques of extending to approved research the same rights and obligations of access enjoyed by officers of the agency need to be considered in terms of better access.*

*12. *There is considerable potential for development of more economical and responsive customized-user services, such as 1) record linkage under the protection of the statistical office, 2) special tabulations, 3) public use samples for special purposes. Such services must often involve some form of cost recovery.*

*13. *Some research and statistical activities require the linking of individual data for research and statistical purposes.*

*14. *Professional organizations or national organizations should have codes of ethics for their disciplines with respect to the utilization of individual data for research and statistical purposes. These ethical codes should furnish mutually agreeable standards of behavior governing relations between providers and users of governmental data.*[4]

*15. *Users of microdata should be required to sign written undertakings for the protection of confidentiality.*

*16. *Considerable efforts should be made to explain to the general public the procedures in force for the protection of the confidentiality of microdata collected and disseminated for research and statistical purposes.*

17. *The right of privacy is evolving rather than static, and closely related to how statistics and research are perceived. Therefore, statisticians and researchers have a responsibility to contribute to policy and legal definitions of privacy.*

18. *Public concern with privacy in the collection and utilization of individual data can be addressed in part as follows:*

    1) *voluntary data collection, whenever practicable*
    2) *advanced general notice to respondents and informed consent, whenever practicable*
    3) *provisions for public knowledge of data uses*
    4) *public education on the distinction between administrative and research uses of information.*

Flaherty summarizes the discussion on each of these principles in a concise form. Appendix I includes the conference participants. Appendix II summarizes the themes which emerged from the sessions on the United Kingdom (public relations, uses of information, and dissemination of data) and Canada (dissemination and uses of data). Appendix II describes the agenda for the General Sessions: goals, disseminating data, regulation of dissemination, forms of microdata dissemination, accountability of researchers, and privacy and data collection. Appendix IV lists the 18 principles agreed upon by the participants.

III.

Data archives are concerned with the preservation of information. Not only are there technical issues related to preservation and release of statistical information, but also ethical, moral and legal considerations about the release of confidential information on individuals. What should be the role of the data archive in accepting information of a confidential nature? Should the archive agree to preserve this information? What sorts of restrictions on access should archives adhere to? If the archive is to distribute confidential information, what sorts of protection should be applied to this information?

These and other questions concerning the maintenance and release of confidential information are questions which seem to have been raised by few data archives and organizations outside the federal statistical agencies such as the Bureau of the Census. Most of the information which the data archive has dealt with has been of an anonymous nature and with few exceptions little confidential data have been deposited with the data archive. [This discussion does not pertain to the National Archives, which has had a long history of concern with this issue and has developed various mechanisms for dealing with the problem.] Nevertheless, confidential information is being collected and archives as such do play a role as preservors of this information. In addition, there are a number of data organizations, integral parts of (survey) research organizations, which are responsible for maintaining the data files created by their researchers; there are data libraries, which although do not have a mission to preserve original data, sometimes (because of their experience with data generation and processing) become involved in projects where confidential data are collected.

At the same time public concern about the release of confidential information by statistical agencies has led to restrictions on access to information needed for research and public policy development by the data archive's clientele.

This suggests that the issue is of immediate and continuing attention by the data archiving community. What to do about confidential data, what is the role of the archive, and what are the archive's responsibilities in this area are difficult questions and have no easy solutions. The problem does suggest however that it would be wise for a data archive to formulate a coherent policy on the preservation and release of statistical information which potentially invades individual privacy.

While I have not suggested any guidelines that a data archive can follow for formulating a policy on the archiving and dissemination of confidential data, I hope that this discussion will stimulate further thoughts by the archive community and a continuing dialogue in this Newsletter.

Footnotes and References

[1] Some introductory materials to the issue of privacy/confidentiality are:

1) The House and Senate hearings held during the middle 1960's, which dealt with the creation of a Federal (National) Data Center, contain a great deal of discussion on the confidentiality problems inherent in a government data center.
    - U.S. Congress. House. Committee on Government Operations. Special Subcommittee on Invasion of Privacy. The Computer and Invasion of Privacy. Hearings, 89th Cong., 2d sess. Washington: G.P.O, 1966. 318 pp. (Appendix 1: "The Ruggles Report," pp. 195-253;
    - U.S. Congress. Senate. Committee on the Judiciary. Subcommittee on Administrative Practice and Procedure. Computer Privacy. Hearings, 90th Cong., 1st sess., on S. Res. 25. Washington: G.P.O., 1967. 388 pp. ("The Kaysen Report," pp. 357-359.)
    - U.S. Congress. House. Committee on Government Operations. Privacy and the National Data Bank Concept. 90th Cong., 2d sess., H. Rept. 1842. Washington: G.P.O., 1968. 34 pp.

2) The Alan Westin book, Privacy and Freedom (New York: Atheneum, 1967), provides a basic introduction to the problem. See also his book Data Banks in a Free Society (New York: Quadrangle Books, 1972).

3) It is useful to consult the various pieces of legislation which have been passed in the United States, such as the Privacy Act of 1974, the Freedom of Information Act, Federal Reports Act, Fair Credit Reporting Act, and the Family Educational Rights and Privacy Act of 1974. Summaries of the contents of these Acts are contained in Kent Greenawalt's monograph, Legal Protections of Privacy, Final Report to the Office of Telecommunications Policy (Executive Office of the President). (Washington, D.C.: Superintendent of Documents, U.S. Government Printing Office).

4) Legislation has been considered and promulgated in other nations. Orjar Oyen, in his article, "Social Research and the Protection of Privacy: A Review of the Norwegian Development" (*Acta Sociologica*: 19 [1976], pp. 249-262), discusses what is occurring in Norway. The impact of the Swedish Data Act is discussed in the proceedings of a symposium on *Personal Integrity and the Need for Data in the Social Sciences*, held at Hässelby Slott, Stockholm, March 15-17, 1976, sponsored by the Swedish Council for Social Science Research. Other discussions have taken place and are reported in proceedings: Stig Stromholm, *Right of Privacy and Rights of the Personality: A Comparative Survey* (Stockholm: P.A. Norstedt and Söners Förlag, 1967), which is a working paper prepared for the Nordic Conference on Privacy organized by the International Commission of Jurists, held in Stockholm, May 1967. See also *The First International Oslo Symposium on Data Banks and Society*, by Universitetsforlaget in Oslo in 1972.

5) For discussions about statistical techniques to protect confidential information, see the Gwendolyn B. Moore *et al*. report, *Accessing Individual Records from Personal Data Files Using Non-Unique Identifiers*, prepared for the Institute for Computer Sciences and Technology of the National Bureau of Standards in Washington, D.C. (NBS Special Publication 500-2, U.S. Department of Commerce, National Bureau of Standards); Appendix A ("Confidentiality-Preserving Modes of Access to Files and to Interfile Exchange for Useful Statistical Analysis" by Donald T. Campbell *et al*.) in *Protecting Individual Privacy in Evaluation Research*, published by the National Academy of Sciences (Washington, D.C., 1975) (included in this monograph is an excellent bibliography); *Proceedings of the Second Midwest Conference on Confidentiality of Health and Social Service Records: Where Law, Ethics, and Clinical Issues Meet* (Chicago: University of Illinois, 1976). Robert Boruch, director of the Evaluation Research Program at Northwestern University, has been very kind to supply me with a number of articles and references to techniques for ensuring confidentiality. Some of the articles by him include "Relations Among Statistical Methods for Assuring Confidentiality of Social Research Data (*Social Science Research* I, 403-414 [1972]); "Strategies for Eliciting and Merging Confidential Social Research Data" (Nejelski, P. [Ed.], *Research in Conflict with Law and Ethics* (Cambridge: Ballinger, 1976); "Statische unde methodische prozeduren zur sickerung der vertraluichkeit bei forschung" (Eser, A. and K.F. Schumann [Eds.], *Forschung im Konflikt mit Recht un Ethik* [Stuttgart: Ferdinand Enke Verlag, 1976]); "Educational Research and the Confidentiality of Data: A Case Study", *Sociology of Education* 44, 1971, 59-85; "Methodological Techniques for Assuring Personal Integrity in Social Research", September 1976 (prepared as background research for Evaluation Research Program's Project on Secondary Analysis.

6) Most of the articles cited above also discuss the ethical issues inherent in the confidentiality problem. In addition to the Boruch article published in the *Sociology of Education*, and the Nejelski (editor) book, also useful is the Boruch and Joseph S. Cecil article, "Is a Promise of Confidentiality Necessary? Sufficient?", which appears as chapter 3 of *Methods for Assuring Confidentiality of Social Research Data*, prepared for the American Psychological Association Task Force on Privacy and Confidentiality and National Academy of Sciences Committee on Federal Statistics, Panel on Privacy and Confidentiality as Factors in Survey Response (Background Research).

[2] David H. Flaherty. "Report of the Bellagio Conference on Privacy, Confidentiality, and the Use of Government Microdata for Research and Statistical Purposes, Lake Como, Italy, August 16-20, 1977." (Draft.) University of Western Ontario, September 17, 1977. (Mimeographed.)

[3] For a brief description of the Privacy Project, see Flaherty's paper, "Privacy and Access to Government Microdata," (revised, November 8, 1976) prepared for the Executive Seminar on "Expanding the Right to Privacy: Research and Legislative Initiatives for the Future," Washington, D.C., October 14-15, 1976, sponsored by the Washington Public Affairs Center, University of Southern California.

[4] The Inter-university Consortium for Political and Social Research (ICPSR) has developed a statement on the archiving, processing, and release of confidential data. I include it because it represents the fullest statement to date which describes what ICPSR processing and dissemination responsiblities are in this area.

PRESERVATION OF CONFIDENTIALITY

The issue of confidentiality arises primarily, but not exclusively, in the case of data collections that include information that is or is seen as potentially damaging or threatening to respondents, or when a promise of confidentiality has been given respondents in the process of data collection, and when information is included that allows or potentially allows identification of individual respondents. The confidentiality issue may appear most pressing in the case of data pertinent to elite populations. Such data are now increasingly available, and elite populations are more visible, are or are seen as at greater risk, and are more easily identifiable by means of a few key variables than are respondents in a mass survey. Also contributing to the gravity of the issue is the sensitivity of the public as a whole, and public figures more particularly, to the "potential" uses of data obtained through personal interviews. The increased research focus on elite populations and the resultant increased availability of data combined with the need to facilitate extended research use of data collections makes it necessary for the Consortium to develop policies and procedures for legitimate data dissemination without jeopardizing the rights of respondents to confidentiality.

It should be noted, of course, that the issue of confidentiality is not confined to information collected through personal interviews nor is it limited to elite populations or even individuals. Data collected through means other than personal interviews--through laboratory experiments, for example--also present the issue, and data collected from public record sources, such as court records, can be seen as unnecessarily and unjustifiably damaging to individuals unless anonymity is provided. Data on organizations or particular groups can also present confidentiality issues as can data relevant to deviant populations and the issue is also raised by mass surveys. While the following policies are intended to apply specifically to individual level data for elite populations, they are also seen as applying in more general terms to other categories of data such as those suggested above.

To protect the identity of respondents, the following guidelines will be applied:

1. The Archive will not accept any documentation, list or data files that explicitly identify respondents by name except in the case of data that are in publicly available sources and which do not present problems of confidentiality.

2) In cooperation and consultation with principal investigator(s) the Consortium staff will assist in the identification of "sensitive" variables but reserves the right to extend the list of variables to be deleted, aggregated or otherwise masked beyond those identified by the principal investigator(s)

3) Variables will be flagged when a) they allow identification of particular respondents, and b) when used in combination with other variables, they also identify particular respondents.

4) The list of sensitive variables will be presented to the principal investigator(s) with recommendations for which variables should be deleted, collapsed or grouped (such as age) or have the code category descriptions changed to generalize the category.

5) A version of the dataset which incorporates the above recommendations will be produced for general distribution.

6) The codebook for the version of the data to be distributed will include documentation for the masked variables, a description of the nature of the deletions and maskings and frequencies for the masked variables.

7) Requests for data reductions (for example, cross-tabulations) or aggregations involving the masked variables will be accepted, but the Consortium staff will judge whether the requested data reductions or aggregations will preclude identification of individuals and will supply only those results that do so. Decisions will be made in consultation with principal investigators in order to insure that protection of confidentiality does not unnecessarily limit researchers' access to analytic information.

8) The original version of the data will be maintained under security (see document entitled "Physical File Security").

9) Public record data will not be added or merged into a survey data file if that data allows identification of individuals. If public record data have been added to the survey data file and identification of individuals as a consequence is possible, these data will be removed from the survey data file and retained separately without a linking variable that would allow the two files to be merged again.

The proposed method for handling confidentiality attempts to protect the respondents without destroying the meaning of the data. If this process renders the data meaningless in any given case, the data could be maintained under security by the Consortium and not be generally distributed. The documentation could be distributed upon request and only data reduction requests would be accepted.

If none of the above is acceptable, the data cannot be archived.

The Data and Program Library Service has attempted to identify donor, archive and user responsibilities in a simpler form, since it involves itself less frequently in processing a data file. Appended is a first (working) version of the DPLS "authorization form" which is agreed upon by the donor when depositing data at DPLS, DPLS when accepting responsibility for a file and the user when receiving a copy of the data. Of particular relevance to the readers are lines #3-5, 10-18, 27-32. As the readers will note, DPLS considers it essential that restrictions may be "relaxed" upon review and approval of data access by the DPLS Faculty Advisory Committee.

DATA AND PROGRAM LIBRARY SERVICE
4451 Social Science Building
University of Wisconsin-Madison
Madison, Wisconsin 53706
608-262-7962

AUTHORIZATION FORM

The Data and Program Library Service serves as an archive for social science related quantitative machine-readable data for the University of Wisconsin-Madison. When authorized as a permanent repository for a data file, the archive assumes responsibility for the condition of the data, changes in the data, distribution, and protection against the release of information deemed confidential. Unless otherwise indicated by the donor of the data, distribution of the file shall be limited to individuals or institutions which shall utilize the data for statistical or research purposes.

The donor of a data file(s) is requested to supply DPLS with information on any restrictions placed upon dissemination of the file. DPLS adheres to these restrictions. DPLS is responsible for notifying the potential user of a file that the file cannot be reproduced without the written consent of the donor. DPLS prefers that a time limit of not more than three (3) years be applied to restrictions on the file, unless confidentiality constraints limit in any way distribution of the data. In that case, unless the archive is assured by the donor of the data and the archive's own evaluation of the file that confidentiality requirements can be met, public access to the file will remain restricted.

DPLS requires that the user of the data agree to the following conditions for receipt of the file:

(a) Neither the transmitted file(s), listing(s), nor any copy thereof in whole or in part, shall be disseminated, or sold, to any further party;

(b) The data provided shall be used solely for professional or other official purposes related to teaching, research, public service, public policy, and planning;

(c) Publication or dissemination of data, or the results of analyses of the data provided, shall be based on a minimum of five individual subjects and/or three organizations per distributional or tabulated cell and neither any individual subject or organization shall be explicitly identified. Where exception to this requirement is requested, permission for access to the data shall be granted by the Data and Computation Center Faculty Policy Committee upon review of the research needs of a particular project;

(d) An abstract or copy of any published document upon which the analysis is based be provided to DPLS;

(e) DPLS will not be responsible for losses or inconvenience resulting from delayed delivery or errors due to defects attributable to the requestor's own equipment. DPLS will bear the cost of replacing the data in the event that the defect is attributable to DPLS's work;

(f) Responsibility for the accuracy of the data and documentation rests with the donor unless DPLS has itself been responsible for producing the file or documentation.

Please sign below if you agree to the above restrictions.

Date ____________________

Donor: Name ____________________
Address ____________________
Telephone # ____________________

User: Name ____________________
Address ____________________
Telephone # ____________________

DPLS ____________________

(1/2)

DISCUSSION PAPER/ donald f. harrison

On Work in Progress: Directory of Directories

Donald F. Harrison
National Archives
Washington, DC

The US Action Group on Process Produced Data Files, in its mandate to prepare a "Directory of Directories", (IASSIST Newsletter Vol. I, No. 3, May 1977), has prepared the following initial entries and desires feedback from the membership. With two exceptions, entries describe only social science data files, printed and available in the United States. Because of the nature of where files are created, plus the specialized knowledge of the list makers, this list is heavy on Federal directories.

The committee seeks information on: 1) additional directories not listed and available for description; and, 2) additional or different information that may be desireable in the entry format. Please send any suggestions to:

Donald F. Harrison
Chairman
Process Produced Data AG
Machine-readable Archives Division (NNR)
National Archives
Washington, DC 20408

Association of Public Data Users. (APDU). Data File Directory. August 1977. 644 pp. typescript. Contains APDU individual and organizational membership directory as well as descriptions of contributing organizations. Lists of selected data files available from each contributing organization with file name and keyword indices. Available to membership only. For membership information, write: APDU, Box 9287 Rosslyn Station, Arlington, VA 22209.

Inter-university Consortium for Political and Social Research. (ICPSR. Guide to Resources and Services. (annual publication). Contains data files available for purchase, technical configurations, prices. Write: ICPSR, University of Michigan, Ann Arbor, MI 48104.

Library of Congress. Machine-Readable Cataloging (MARC); Price Announcements and Subscription Rates. Washington, DC, undated, 7 pp. typescript. Describes 17 subscription service data bases in MARC II. Available from L/C. Prices, available technical configurations. Write: Library of Congress, Processing Department, Cataloging Distribution Service, Washington, DC 20540.

Sessions, Vivian S. (ed.) Directory of Data Bases in the Social and Behavioral Sciences. New York: Social Associates/International Inc. 1974. hb 300 pp. Badly out of date. Describes files and other data configurations at 685 institutions/public and private, US and foreign. Describes at the file level, availability, etc.

Texas Christian University, Drug Abuse Epidemiology Data Center (DAEDAC) Newsletter. Describes files available from DAEDAC, including the "Original Data File Catalog". For further information, write: Associate Director, Dr. LaVerne Knezek, DAEDAC, Texas Christian University, Fort Worth, TX 76129.

U.S. Bureau of the Census. Catalog. (issued quarterly). Contains social and economic data compiled by the Bureau on a quarterly basis. Write: Users' Service Staff, Data User Services Office, Bureau of the Census, Washington, DC 20233.

U.S. Department of Commerce. National Technical Information Service. Directory of Computerized Data Files and Related Software. (published periodically). Contains data files available from NTIS or from Federal agencies. Technical configurations, prices. Write: NTIS, 5285 Port Royal Road, Springfield, VA 22161.

U.S. Department of Health, Education and Welfare. Office of Education. Directory of Federal Agency Education Data Tapes. By Barbara A. Feller. (Pub. NCES 76-206). Washington: GPO 1976. 177 pp. apps. describes educational data bases available to the general public from Federal agencies. Prices, technical configurations. Available GPO.

U.S. Department of Health, Education and Welfare. Public Health Service. Directory of Automatic Data Processing Systems in the Public Health Service. (Published annually). 510 pp. typescript with indices. Describes approx. 500 files in PHS. Directory not yet available to public.

U.S. Department of Health, Education and Welfare. Social Security Administration. Some Statistical Research Resources Available at the Social Security Administration. A description of the available lifetime overall earnings data files from the several-purpose research files making up the continuous Work History Sample (CWHS).

U.S. Department of Health, Education and Welfare. Standardized Micro-Data Tape Transcripts. (DHEW Pub. No. 76-1213). Washington: GPO 1976. 30 pp. Describes data bases for health statistics available from the Federal government. Prices, technical configurations. Available from GPO.

U.S. Department of Labor. Bureau of Labor Statistics. BLS Data Bank Files and Statistical Routines. (draft brochure). Describes 30 data files available from the Bureau of Labor Statistics. Describes publications. Write: Commissioner, BLS, Dept. of Labor, Washington, DC 20212.

U.S. General Accounting Office. 1976 Congressional Sourcebook Series (OPA 76-23) Federal Information Sources and Systems; A Directory for the Congress. Washington, DC: GPO 1976. 456 pp. paperback. An indexed reference guide to over 1,000 Federal sources and information systems in 63 Federal agencies, which contain budgetary, fiscal and program-related data. Available from GPO.

U.S. National Archives and Records Service. Catalog of Machine-Readable Records in the National Archives of the United States. Washington: GPO, 1977. 37 pp. Describes 99 files created by Federal agencies and retained in the National Archives. Prices, available technical configurations. Write: National Archives (NNR), Washington, DC 20408.

University of Waterloo. Leisure Studies Data Bank. April, 1977. 23 pp. Lists 24 files created by private and Federal Canadian agencies. In French and English. Write: Dept. of Recreation, University of Waterloo, Waterloo, Ontario, Canada. N2L 3G1.

DISCUSSION PAPER / gary m. grandon

Using SPSS MULT RESPONSE to Generate Filtered Marginals

Gary M. Grandon
Social Science Data Center
The University of Connecticut, Storrs

With the release of SPSS's Version 7.0 this year, several additions have been added to its already extensive battery of analysis programs. MULT RESPONSE is one of these programs. It generates frequency counts and bivariate tables for "dummy" coded multiple response questions. These multiple response variables are a real nuisance to the analyst and the program at face value provides a means for their interpretation. Considerations for the analysis of multiple response data are not the subject of this paper though, but rather the use of this program for quite another purpose; the generation of "filtered marginals."

Filtered marginals are typically basic frequency counts and percentages for specified variables for each of a number of subpopulations within a study. SPSS has in the past provided *SELECT IF statements to facilitate the processing of subpopulations. Each such statement followed by FREQUENCIES procedure statements will generate filtered marginals. The inherent problem with this type of coding is that each *SELECT IF statement requires a complete reading of the data set. For large studies this procedure can be quite inefficient! Another alternative to the *SELECT IF procedure would be to create a new variable with codes indicating each filter group (subpopulation). Then the very fast CROSSTABS procedure can be used with the new filter group variable as one variable in the bivariate table and the variable upon which frequencies are to be generated as the other. A large number of frequencies can be generated in this manner with only one read of the data. The problem with this CROSSTABS procedure is that it requires that each filter group or subpopulation be discrete. That is, any one subject can only be in one filter group. This restriction is removed by using the MULT RESPONSE technique suggest in this paper. The following setup using the SPSS - OSIRIS interface to shorten the presentation illustrates the use of the MULT RESPONSE procedures to perform filtered marginals:

```
1              16

RUN NAME       YANKELOVICH---COMBINED STUDIES-1976
OSIRIS VARS    V7,V8,V11,V13,V18,V28,V64,V75
INPUT MEDIUM   TAPE
N OF CASES     UNKNOWN
WEIGHT         V75
IF             ((V7 EQ 1 OR 2 OR 3) AND (V8 EQ 6) AND (V11 EQ 1) AND (V18 EQ 1)F1=1
IF             ((V7 EQ 5 or 6) AND (V8 EQ 1 OR 2) AND (V11 EQ )3) AND (V18 EQ 1)F2=1
MULT RESPONSE  GROUPS=FILTERS (F1 F2(1))/
               VARIABLES=V13(1,4),V28(1,3),V64(1,3)/
               TABLES=FILTERS BY V13 TO V64/
STATISTICS     1
READ INPUT DATA
FINISH
```

Not only does the technique illustrated above produce the same kind of results as the *SELECT IF process but it formats the output in such a way as to allow the user to scan differences across subpopulations within a single tabular display. A single procedure statement like the one above can process 20 filter groups and up to 100 analysis variables. Each frequency variable will have its own table.

A resource expenditure comparison between the *SELECT IF procedure and the MULT RESPONSE procedure was performed on the 1976 Yankelovich Combined Study archived at The Social Science Data Center of the University of Connecticut. The study has a N of 7,977 subjects. Nineteen filter groups were identified and frequencies for 45 analysis variables were obtained using both techniques. Table 1 provides relevant resource comparisons for the techniques. The CROSSTABS technique was not included due to its unique group membership constraint. Clearly costs are much lower for the MULT RESPONSE procedure.

Table 1. A resource comparison* of the *SELECT IF and MULT RESPONSE techniques for filtered marginals.

RESOURCE	*SELECT IF	MULT RESPONSE
NUMBER OF JOBS	2	1
KBYT-SEC	21,016	24,957
ELAP-KBS	180,493	73,752
READER EXCP**	93	31
PRINTER EXCP	14,630	3,844
PUBLIC DISK EXCP	10,476	360
PRIVATE DISK EXCP	10	5
TAPE EXCP	72	36
TAPE MOUNTS	2	1
COST	$31.70	$13.39
ELAPSED TIME	13.48 min.	6.23 min.

*Comparisons were performed on the University of Connecticut's Research Computer Center IBM 360/65 - IBM 370/155 systems running under OS and shared spool HASP. Local costs may vary from installation to installation.

**Each EXCP represents the "Execution of a channel program" and indicates the movement of one block of data.

The MULT RESPONSE procedure can further be extended to filtered bivariate tables by including a second BY statement on the TABLES card. This technique is not illustrated due to space limitations. Readers are referred to their local installations for further documentation of SPSS Version 7.0.

DISCUSSION PAPER/richard c. roistacher

The following article describes the work of Richard C. Roistacher and Barbara Noble at the Center for Advanced Computation at the University of Illinois. They are involved in the development of guidelines for the descriptive materials which accompany a data file.

A Source Documentation Style Manual

by
Richard Roistacher
Center for Advanced Computation
University of Illinois
Urbana, Illinois

Barbara Noble and Richard Roistacher of the University of Illinois' Center for Advanced Computation are currently developing a style manual for the documentation of machine readable data. The manual, which is being developed as part of a project funded by the U.S. Department of Justice's Law Enforcement Assistance Administration, is presently available in draft form. The manual, conforming to the naming conventions of the U.S. Government's Federal Information Processing Standards, is titled LEAA Research Support Center Machine Readable Source Documentation System: User's Guide.

The style manual is designed to serve the needs of data producers, archivists, and users. It was developed from several current series of documentation, as well as from the LEAA project's experience in archiving a number of large criminal justice data files.

The manual gives an annotated example of each major section of the source documentation for a machine readable data file. The manual's major sections describe the cover, abstract, introduction, codebook, and appendices of source documentation. The section on the cover gives advice on the titling, authorship, and citation of files and documentation. The manual's chapters on the abstract and introduction sections gives a form for abstracts, and gives advice on how to describe the file's collection, methodological, and processing history. The chapter on the codebook section gives formats for data item blocks for several types of data. The chapter on appendices gives formats for listings of known errors, definitions of terms, dictionary descriptions, description of data shipments, and extended code listings.

At present, the examples in the manual cover only rectangular files of survey data. Additions to the manual will cover the documentation of hierarchical and other complex files, time series, network data, and generalized arrays.

The documentation manual is designed to be independent of any particular manual or computer based system for producing documentation. However, the LEAA project at the Center for Advanced Computation has developed a computer system for producing machine readable source documentation conforming to the manual. The system merges the information in an OSIRIS III data dictionary file into a file of documentation text. The documentation text is processed with the University of British Columbia's FMT document processing program to yield a finished codebook, which can be printed or copied to tape. The system will automatically reformat and reorder the documentation to match a data file which has been subsetted, reordered, or merged. New page numbers, tables of contents, cross references, and indices are generated by the FORMAT program without any intervention by the user. A later version of the system, which runs on IBM hardware, will extract format information from SPSS system files.

This winter, Noble and Roistacher will hold a conference on the documentation of machine readable data. One of the major purposes of the conference will be to expand and refine the current draft of the documentation style manual.

The authors would welcome advice and comments from data producers, archivists, and users. Copies of the current draft are available to interested people, and any comments would be welcome. Copies are available from Barbara B. Noble or Richard C. Roistacher, Center for Advanced Computation, University of Illinois, Urbana, IL 61801.

[Editor's Note: The meeting was held on October 27 and 28, 1977 in Boston. Revisions to the Manual were recommended. A full report will appear in the next issues of IASSIST as Roistacher and Noble further refine the Manual.]

iassist NEWSLETTER vol. 1, no. 4

DATA ORGANIZATION REGISTRY FORM
june1977

Data Organization Registry Form

Version 1.0 June 1977

This form concerns official policies and procedures of DATA ORGANIZATIONS. Respondents should complete the form on behalf of their Data ORGANIZATION. ('DO' refers to the DATA ORGANIZATION.)

This form was developed at the Western European IASSIST workshop in Copenhagen June 27-29, 1977.

Table of Contents

1.0 Identification 1
2.0 Origin and History of Data Organization 3
3.0 Affiliations 4
4.0 Structure 5
   4.1 Funding 5
   4.2 Service Community 5
   4.3 Personnel 6
   4.4 Computer Use 7
5.0 Functions of the Data Organization 10
   5.1 Publications 10
   5.2 Products and Services 12
   5.3 Training Courses 13
6.0 Data Holdings 15
Acknowledgements 20

1.0 Identification
1.1.1 Date of completion of this form: ______
1.1.2 Name of respondent: ______
1.1.3 Position of respondent in DO: ______
1.2.1 Official name of the DO: ______
1.2.2 Acronym for official name: ______
1.2.3 Mailing address of the DO:
1.2.3.1 (Street) ______
1.2.3.2 (City) ______
1.2.3.3 (Country) ______
1.2.3.4 (Postal code) ______
1.2.3.5 (Cable address) ______
1.2.4.1 Central telephone number for the DO:
1.2.4.2 Country code: ______
1.2.4.3 Area code: ______
1.2.4.4 Subscriber number: ______
1.2.4.5 Telex number: ______
1.3.1 Name of parent institution/organization of DO (if the DO has no parent organization, please write NONE): ______
1.3.2 Indicate how the parent institution/organization is classified:
1.3.2.1 None ______
1.3.2.2 Academic institution ______
1.3.2.3 Private non-profit organization ______
1.3.2.4 Commerical (profit-making) organization ______
1.3.2.5 National government ______
1.3.2.6 Regional government (incl. provincial/state) ______
1.3.2.7 Local government ______
1.4 If the DO has no parent institution/organization, please indicate how you classify the DO with respect to the responses in 1.3.2. If the DO has a parent institution/organization, write 'not applicable'. ______
1.5 The year the DO was founded: ______
2.0 Origin and History of Data Organization
Please provide a short paragraph which includes the most salient information.
______

3.0 Affiliations
Indicate the principal organizations with which the DO is affiliated:
3.1 IFDO ______
3.2 CESSDA ______
3.3 ECPR ______
3.4 ICPSR ______
3.5 Roper Center ______
3.6 Other (specify) ______
3.7 ______
3.8 ______
3.9 ______
4.0 Structure
4.1 Funding
Please indicate the percentage of funds received during the past year from the following sources:
4.1.1 University ______
4.1.2 Private/non-profit organizations (incl. foundations) ______
4.1.3 Commercial (profit-making) organizations ______
4.1.4 National government (incl. Research Councils) ______
4.1.5 Regional government (incl. provincial or state) ______
4.1.6 Local government ______
4.1.7 Other (specify): ______ ______
Total 100%
4.2 Service Community
Please indicate the percentage of service provided to each of the following during the past year:
4.2.1 Your own organization ______
4.2.2 Your parent organization community ______
In addition:
4.2.3 Academic communities ______
4.2.4 Private/non-profit organizations ______
4.2.5 Commercial (profit-making) organizations ______
4.2.6 International community ______
4.2.7 National government organizations ______
4.2.8 Regional government organizations ______
4.2.9 Local government organizations ______
Total 100%
4.3 Personnel
4.3.1 Complement of full-time staff in man-years
4.3.1.1 Number of administrative/managerial positions in man-years ______
4.3.1.2 Number of secretarial/clerical positions in man-years ______
4.3.1.3 Number of data acquisition positions in man-years ______
4.3.1.4 Number of file development positions in man-years ______
4.3.1.5 Number of user services positions in man-years ______
4.3.1.6 Number of programmers in man-years ______
4.3.1.7 Number of terminal operators/keypunchers in man-years ______
4.3.1.8 Other positions in man-years (specify) ______ ______
4.3.2.1 Number of full-time employees ______
4.3.2.2 Total number of part-time employees ______
4.3.2.3 Number of part-time employees holding joint appointments with parent organization (i.e. teaching positions, etc.) ______
4.4 Computer use
4.4.1 Indicate how the DO pays for computer use:
4.4.1.1 Budgeted funds
4.4.1.2 Subsidy ______
4.4.1.3 Combination of budgeted funds and subsidy ______

4.4.2 Please list the types of computers available to the DO and their ownership:

	Manufacturer and model	owner: DO	owner: parent org.	owner: other
4.4.2.1	________	____	____	____
4.4.2.2	________	____	____	____
4.4.2.3	________	____	____	____
4.4.2.4	________	____	____	____
4.4.2.5	________	____	____	____

1.4.3 Can the DO process multipunched data? yes ____ no ____

4.4.4 Indicate which of the following items the DD has access to:
4.4.4.1 Nine track tape drives ____
4.4.4.2 Seven track tape drives ____
4.4.4.3 6250 BPI density ____
4.4.4.4 3200 BPI density ____
4.4.4.5 1600 BPI density ____
4.4.4.6 800 BPI density ____
4.4.4.7 556 BPI density ____
4.4.4.8 Cassette tapes ____

4.4.5 Indicate which of the following equipment the DD has access to:
4.4.5.1 Magnetic card reader ____
4.4.5.2 Optical character recognition device ____
4.4.5.3 Paper tape reader ____
4.4.5.4 Dismountable discs ____

4.4.6 Indicate the number of the following devices that the DO has access to in its offices or elsewhere, and who owns these devices:

		where (number)		ownership (number)	
		DO offices	elsewhere	DO	other
4.4.6.1	Terminals	____	____	____	____
4.4.6.2	Card reader	____	____	____	____
4.4.6.3	Key punch	____	____	____	____
4.4.6.4	Upper case line printer	____	____	____	____
4.4.6.5	Upper and lower case line printer	____	____	____	____
4.4.6.6	Counter-sorter	____	____	____	____
4.4.6.7	Card reproducer	____	____	____	____
4.4.6.8	Paper tape reader	____	____	____	____
4.4.6.9	Magnetic card producer	____	____	____	____

4.4.7 Indicate additional hardware which the DO has access to (specify):
4.4.7.1 ________
4.4.7.2 ________
4.4.7.3 ________
4.4.7.4 ________
4.4.7.5 ________

5.0 Functions of the Data Organization

5.1 Publications

5.1.1 Does the DO publish a newsletter? yes ____ no ____
5.1.1.1 If a newsletter is published:
5.1.1.2 Name of newsletter: ________
5.1.1.3 Frequency: ________
5.1.1.4.1 Languages of publication: ________
5.1.1.4.2 ________
5.1.1.4.3 ________

5.1.2 Does the DO publish a catalogue of its holdings? yes ____ no ____
5.1.2.1 If a catalogue is published:
5.1.2.2 Frequency of publication: ________
5.1.2.3.1 Languages of publication: ________
5.1.2.3.2 ________
5.1.2.3.3 ________

5.1.2.4 Give a bibliographic reference for the most recent version of the DO's catalogue of holdings:
5.1.2.4.1 Author: ________
5.1.2.4.2 Title: ________
5.1.2.4.3 Date of publication: ________
5.1.2.4.4 Publisher: ________
5.1.2.4.5 Place of publication: ________

5.1.3.1 Does the DO publish an information brochure:
5.1.3.1.1 For users ____
5.1.3.1.2 For depositors ____
5.1.3.1.3 Combined brochure for users and depositors ____

5.1.4 If the DO publishes other types of documents, such as monographs, research reports, annual reports, research catalogues, programmer manuals, and the like, please give bibliographic references for the 5 most recent publications:
5.1.4.1.1 Author: ________
5.1.4.1.2 Title: ________
5.1.4.1.3 Date of publication: ________
5.1.4.1.4 Publisher: ________
5.1.4.1.5 Place of publication: ________
5.1.4.2.1 Author: ________
5.1.4.2.2 Title: ________
5.1.4.2.3 Date of publication: ________
5.1.4.2.4 Publisher: ________
5.1.4.2.5 Place of publication: ________
5.1.4.3.1 Author: ________
5.1.4.3.2 Title: ________
5.1.4.3.3 Date of publication: ________
5.1.4.3.4 Publisher: ________
5.1.4.3.5 Place of publication: ________
5.1.4.4.1 Author: ________
5.1.4.4.2 Title: ________
5.1.4.4.3 Date of publication: ________
5.1.4.4.4 Publisher: ________
5.1.4.4.5 Place of publication: ________
5.1.4.5.1 Author: ________
5.1.4.5.2 Title: ________
5.1.4.5.3 Date of publication: ________
5.1.4.5.4 Publisher: ________
5.1.4.5.5 Place of publication: ________

5.2 Products and services

5.2.1 Indicate which of the following products or services are offered by the DO:
5.2.1.1 Abstract writing ____
5.2.1.2 Cataloging ____
5.2.1.3 Data cleaning ____
5.2.1.4 Data documenting ____
5.2.1.5 Data entry ____
5.2.1.6 Data file copying ____
5.2.1.7 Data provision for secondary analysis ____
5.2.1.8 Field work ____
5.2.1.9 Graphics ____
5.2.1.10 Indexing ____
5.2.1.11 Instrument construction ____
5.2.1.12 Interpretive analysis and review ____
5.2.1.13 Reference service ____
5.2.1.14 Referral service ____
5.2.1.15 Research design consultation ____
5.2.1.16 Research contracts (conduct of entire project) ____

5.2.1.17 Statistical analysis ______
5.2.1.18 Tabulations ______
5.2.1.19 Training programmes ______
5.2.1.20 Other (specify): ______

5.3 Training courses

5.3.1 In the last 3 years has the DO offered training courses or workshops? yes ____ no ____

5.3.2 If yes, indicate who were the constituents:
5.3.2.1 Staff of the DO
5.3.2.2.1 Parent organization: faculty ______
5.3.2.2.2 students ______
5.3.2.2.3 staff members ______
5.3.2.3.1 Other organizations: faculty ______
5.3.2.3.2 students ______
5.3.2.3.3 staff members ______

5.3.3 Indicate which of the following courses were offered, and the number of courses offered during the past 3 years:
5.3.3.1 Computer related instruction ______
5.3.3.2 Data analysis ______
5.3.3.3 Data management ______
5.3.3.4 Issue related courses (i.e., social mobility, environmental studies), specify: ______
5.3.3.5 Programme packages:
5.3.3.5.1 graphics ______
5.3.3.5.2 simulation and modeling ______
5.3.3.5.3 statistical ______
5.3.3.5.4 text editors and formatters ______
5.3.3.5.5 other (specify) ______
5.3.3.6 Research methods ______
5.3.3.7 Resources and services of the DO ______
5.3.3.8 Other (specify) ______

6.0 Data Holdings

6.1 Does the DO have a formal method for monitoring on-going research likely to generate data eligible for acquisition? yes ____ no ____

6.1.1 If the DO has a formal monitoring method, please describe this method: ______

6.2 Does the DO have formal deposit arrangements with any of the following:
6.2.1 Granting organizations ______
6.2.2 Government organizations ______
6.2.3 Academic organizations ______
6.2.4 Private/non-profit organizations ______
6.2.5 Commercial organizations ______
6.2.6 Other (specify) ______

6.3 What is the DO's categories for describing the readiness of its archival files for analysis? Please list and define each category and indicate the percentage of files in the DO holdings in each category.

category	definition	approx. % of holdings
____	______	______
____	______	______
	Total	100%

6.4 Does the DO retain a file of data in the condition in which it was received from the depositor? yes ____ no ____

6.5 Does the DO request original unedited (raw) data from the depositor and retain this data in a file? yes ____ no ____

6.6 Indicate the codes and definitions for access categories of the DO holdings, if there are ACCESS RESTRICTIONS imposed on a file:

code	definitions	approx. % of holdings
____	NO RESTRICTIONS	______
____	______	______
	Total	100%

6.7 For data sets most readily available for secondary analysis, indicate if the following documents exist:

	in print	machine readable	computerized information system
Library information	____	____	____
Abstracts	____	____	____
Study description	____	____	____
List of variables	____	____	____
Codebook	____	____	____
Special publications	____	____	____

6.8 For each of the following documents indicate the language(s) in which it is available (eg. English, German, etc.):

	language(s)
Library information	______
Abstracts	______
Study Description	______
List of variables	______
Codebook	______
Special publications	______

6.9 Types of holdings
Indicate the approximate number of holdings in each category and the DO's anticipated interest:

Data type	now hold (number)	would accept additional (check)	anticipated increase next 12 months (number)
Biographic	____	____	____
Census and aggregate	____	____	____
Computer programmes	____	____	____
Event and incident	____	____	____
Facility	____	____	____
Opinion and attitude	____	____	____
Organization/institution	____	____	____
Physical ecology	____	____	____
Survey	____	____	____
Textual	____	____	____
Teaching packages	____	____	____
Test	____	____	____

6.10 How does the DO classify the bulk of its holdings with respect to subject matter fields, especially survey data? (Examples: political studies, leisure studies, history, management studies, international affairs, etc.) ______

6.11 Indicate the geographic scope of the DO's current and anticipated holdings:

	current (check)	anticipated (check)
Own country:		
Local	____	____
Metropolitan	____	____
Regional (province/state)	____	____
National	____	____
Other countries:		
Local	____	____
Metropolitan	____	____
Regional (province/state)	____	____
National	____	____
Cross-national	____	____
International	____	____

iassist NEWSLETTER vol. 1, no. 4

BOOK NOTICES / kathleen m. heim

Machine-Readable Social Science Data - Special issue of the Drexel Library Quarterly. Volume 13 (January 1977). Edited by Howard D. White. Issue available for $5.00 by writing: Drexel Library Quarterly, Graduate School of Library Science, Drexel University, Philadelphia, Pennsylvania 19104.

Contents: "Numeric Data Files: An Introduction." Howard D. White.
"The Pre-Acquisition Process: A Strategy for Locating and Acquiring Machine-Readable Data." Alice Robbin.
"Stalking the Wild Data Set: The Acquisition of Machine-Readable Social Science Data at Home and Abroad." David Nasatir.
"Cataloging Machine-Readable Data Files - A First Step?" Sue A. Dodd.
"Social Science Data Files, the Research Library and the Computing Center." Douglas Ferguson.
"The Information Transfer Process on the University Campus: The Case for Public Use Files." Lorraine Borman and Richard Hay, Jr.
"Training the Professional Data Librarian." Judith S. Rowe and Carolyn L. Geda.

"This is a contribution to a growing discussion between and within the computing and library worlds." - Douglas Ferguson

Though Ferguson is referring to his own article in this special number of the Drexel Library Quarterly, his statement quite accurately describes the entire issue, for this marks the first time that the library press has devoted a substantial amount of space to treatment of the problems that confront the data archivist. The contributors are familiar names to those active in the data archive movement--all have published extensively and have been involved in the conceptual definitions of the principles of data archiving.

Because the intended audience of this issue is the broad range of library and information professionals, the treatment of matters familiar to the working data archivist may seem elementary. Howard D. White's introduction to numeric data files is a careful building of concepts and definitions that provides the newcomer with vocabulary and a general framework on which to base an understanding of the technical articles that follow. For the novice, Douglas Ferguson's article follows White's most logically for he assesses the role of the library in data archive development and provides the rationale for cooperation. His use of the specific example of the Stanford University experience provides a cogent model of library-archive interaction. Lorraine Borman and Richard Hay, Jr. extend the concepts defined by Ferguson by presenting the present-day situation in universities for the use of public files. They attack the problem of low use by focusing upon the kind of public use files in a social science file collection; illustrate how these files can be accessed and manipulated by both novice and sophisticated users; and show how this type of use is related to the utilization of public information files and information systems, and assists in the total information transfer process.

From these articles the more specialized essays by Alice Robbin, David Nasatir, and Sue A. Dodd may be approached. Robbin's scrutinization of the pre-acquisition process is a detailed, logical analysis of the complex factors that enter into the decision to obtain a particular data set. After a general introduction which places the data archive squarely as a type of special library, Robbin explores the five factors which influence the pre-acquisition process and the search strategy for locating data. A diagrammatic presentation of the search strategy is presented with each part carefully explained in an appendix. Nasatir's brief piece, "Stalking the Wild Data Set," is by far the most dramatic contribution to the entire issue. It is a simple, yet vivid rendering of the sometimes Machiavellian, sometimes mendicant contortions which a data archivist must endure to procure material. Working data archivists will empathize, and those unfamiliar with the field will marvel, at Nasatir's travails.

Dodd's article on cataloging Machine-Readable Data Files (MRDF) may have been the most difficult piece to compose. It has fallen to Dodd to unravel the lengthy history of both the library and data archive profession's attempts to establish conformity in the cataloging of MRDF. The article is loaded with acronyms, committees and subcommittees of the ALA, ASIS, IASSIST and AECT. Though difficult at the outset her essay provides a solid introduction to the monumental problems involved in hammering out a cataloging format agreeable to the various factions. She examines progress to the present focusing on the IASSIST Classification Group project designed to test the feasibility of cataloging social science data files according to the ALA subcommittee's recommended procedures. An appendix of sample catalog cards from those participating in the project is included.

Finally, Judith S. Rowe and Carolyn L. Geda discuss the training of the professional data librarian. The profession is still in its formative stages and current practioners are drawn from a diverse group: programmers, social scientists or librarians. The only formal training available is via the Inter-University Consortium for Political and Social Research summer program held at Ann Arbor, Michigan, which is described in detail.

As stated above, the intended audience of this special issue is the working librarian or information professional. This audience requires definitions that may be elementary to the data archivist. However, Howard D. White's editing has produced a diverse publication that operates on several levels. His contributors have at once provided a sourcebook of articles comprehensible to one unfamiliar with data archiving and yet valuable to the experienced archivist. Dodd's patient unraveling of the machinations of cataloging MRDF and Robbin's precise articulation of pre-acquisition procedures are the only explanations of these processes in the literature. The essays devoted to the reference and information component of data archiving extend current ideas about service and begin to outline the emergent user-models.

This issue will stand as a vital resource for data archiving until some more comprehensive publication supersedes it. The only glaring defect is the lack of a simple introduction that provides a framework for the whole. Nevertheless, this issue of the Drexel Library Quarterly is a necessary item for all data archivists. It will provide the profession with a stronger realization of its role in the information transfer process and it will provide traditional librarians with an understanding of the role of data archives in that process.

Richard C. Roistacher and Barbara C. Noble. Computer Network Support of Social Research Communities. University of Illinois. Graduate College. Center for Advanced Computation. Urbana, Illinois 61801.

In March of 1977 following the First IASSIST North American Conference, Richard Roistacher provided participants with publications and technical reports which fell into three major classes: 1) LEAA Data Archive Material; 2) Data Archiving and Analysis Methods; 3) Sociology of Computing. All of the papers advance the conceptualization of the role of data archives in the research community. Mr. Roistacher has agreed to provide copies to interested data archivists who contact him at the Center for Advanced Computation.

In this notice we would like to focus upon one of the papers which falls into the category of the Sociology of Computing: Computer Network Support of Social Research Communities. This paper, a pre-print of an article to be published by Human Organization, sets forth the idea that present computer technology will allow geographically separated social researchers to share data and data analysis tools, collaborate on research and writing, and communicate with their colleagues much as if they were in a single department. A description of the Law Enforcement Assistance Administration (LEAA) computer network based research support center which operates over a commercial communication network is described in detail.

The LEAA data archive and research facility for social scientists and program evaluation groups which has been developed by the Center for Advanced Computation at the University of Illinois to provide research support services to clients scattered around the country is viewed by Roistacher as a prototype. The services offered by a computer network fall into three major categories: data management and archiving; computer and statistical consulting; and communications and documentation.

Roistacher uses the generic term, Research Support Facility (RSF), to describe the multi-purpose service which this prototype provides. Services fall into two modes: those which do not require a computer network connection to the client (data archives; consulting on analysis of archive files; consulting on client-collected data; agency for other archives; bibliographic information retrieval; and software distribution) and those which can only be provided over a direct network connection (shared archival data; shared software; on-line consulting; shared user data; remote technical services; network mail; network conferencing; facilitating the formation of working groups; and a network journal). All services are described in detail. Potential clients for RSF services include academic researchers, independent research organizations, state planning agencies, and program evaluation groups.

Computer and network costs are compared favorably with the costs of using local facilities. Though charges seem high Roistacher demonstrates that they are much lower than travel, postage, typing, duplicating and time considerations under the present system. Management considerations in remote services are treated. These include training users, administrative procedures, accomodation of clients in distant time zones, terminal acquisitions, and clear and self-sufficient documentation.

The social consequences of networking as perceived by Roistacher's observations of the Department of Defense's ARPAnet are discussed and the implications for collaboration among colleagues regardless of geographical considerations explored.

Though data archives are only one component in the larger RSF, they are a vital component and data archivists will want to ponder Roistacher's predictions for computer-based research support. His experiences with a highly developed facility provide a model for future developments in behavioral science communication and augur a new perhaps more democratic era for the invisible college.

Statistical Policy Division. Office of Management and Budget. In Cooperation with the Federal Statistical Agencies with Responsibility for the Collection, Processing, Analysis and Dissemination from Major Federal Statistical Programs. Framework for Planning U.S. Federal Statistics 1978-1989. (Reprinted from the Statistical Reporter, 1974-1976). Copies available from the Statistical Policy Division, Office of Management and Budget, 726 Jackson Place, N.W., Washington, D.C. 20503.

This collection of reprints from the Statistical Reporter details the efforts of the Statistical Policy Division of the Office of Management and Budget to provide for a more systematic review of the needs for improved statistical planning and coordination. A topical outline of the planning framework is presented along with articles on its nature. The goal is improved coordination of the highly decentralized U.S. Federal Statistical System and setting of statistical priorities.

Of particular interest is the article, "User Access-Data Banks," which includes a bibliography of major agency guides and indexes to statistical data. Much concern is expressed over the failure of government agencies to comply with government regulations which provide for improvement of the dissemination of statistical information on a timely basis.

Feedback is requested from users of government generated data. IASSIST members can obtain a copy of the Framework for Planning U.S. Federal Statistics 1978-1989 from the above address and reply to the same on matters concerning shortcomings in the statistical services of the U.S.

NEW ORGANIZATIONS/ REORGANIZATION

UNISIST

At the 19th session of the UNESCO General Conference, a resolution was adopted concerning the future development of the UNISIST program. The UNISIST Newsletter (Volume 4, No. 4, 1976) notes that the "Conference agreed to set up a General Information Programme concerning the activities of the organization in the field of scientific and technical information and of documentation, libraries and archives." Part 5 of the Resolution particularly concerns the data archive community:

> It "authorizes the Director-General to facilitate the implementation of the General Information Programme, by seeing that activities are integrated with a view to: ... (c) contributing to the development of information infrastructures and to the application of modern techniques of data collection, processing, transfer and reproduction, [and] (d) promoting the training and education of information specialists and information users, with particular attention to the needs of the developing countries, especially the problems of transfer of information and data from the technologically advanced countries to the developing nations."

The UNISIST Newsletter issues which were sent to the IASSIST Newsletter Editor contain a great deal of information on upcoming meetings throughout the world which will deal with an assortment of information problems (such as transfer, documentation and systems development, bibliographic data bases, management, training of staff, resource sharing) and descriptions of information systems (such as POPINS [International Population Information System], ISONET [International Standards Information Network]).

IASSIST members in the U.S. may request back copies of the UNISIST/General Information Program Newsletter and ask to be placed on the mailing list by writing to: CISTIP Secretariat, National Academy of Sciences, 2101 Constitution Avenue, N.W., Washington, D.C. 20418. Outside the U.S., IASSIST members should contact their respective national science academy or the UNESCO Division of Scientific and Technological Information and Documentation, 7, Place de Fontenoy, 75700 Paris, France. CISTIP stands for the U.S. Committee for UNISIST. [Committees with the same purpose must have been established in other countries. If IASSIST members have information about this, please pass it on through the Newsletter.]

UNESCO has also created an Ad Hoc Committee on Social Science Information, which will hold its first meeting in Paris in November 1977. It is intended that this Ad Hoc Committee will provide intellectual and policy guidance for the inclusion of social science information in UNISIST. [From a letter by Fred W. Riggs to Judith S. Rowe, 4 October 1977, pp. 1,2.]

INSSIST, the International Social Science Information Services and Technologies, is an action group which intends to serve as a catalyst for the tasks of liaison. INSSIST will meet to exchange information during the forthcoming ISA [?] conference in Washington (Sheraton Park Hotel, February 22-26, 1978).

The director of the UNESCO Division responsible for promoting the international development of social science, Vladimir Mshvenieradze, is expected to attend and report on the work of his division, including the new Ad Hoc Committee. The time and place of the meeting will be announced in the final ISA program. [Announcement of INSSIST meeting from Fred Riggs to IASSIST Editor, received 2 November 1977.]

INTERNATIONAL ASSOCIATION FOR STATISTICAL COMPUTING

E. Lunenberg, director of the Permanent Office of the International Statistical Institute, has sent Wm. Gammell (US Coordinator of DOM AG) the following announcement of the formation of an International Association for Statistical Computing, a section of the International Statistical Institute. IASC will be open to membership among those interested in promoting the theory, methods, and practice of statistical computing. The objectives of the Association will be to foster interest and knowledge in effective statistical computing through international contacts among statisticians, computing professionals, organizations, institutions, governments and the general public in different countries of the world. Special attention will be given to the developing countries. It intends to promote collaborative efforts with international, national, regional and other organizations and institutions with similar aims; foster evaluations of statistical computing techniques and programs; and, facilitate the exchange of computer programs and their documentation. The Association plans programs and meetings particularly in conjunction with sessions of the International Statistical Institute. The organization will formally come into existence and have its inaugural meeting in December 1977, during the 41st Session of the International Statistical Institutes in New Delhi, India.

The program planned for the IASC meeting includes (1) statistical computing and future directions in the context of an international association for statistical computing; a constitutional/business meeting; and, interactions between statistics and computing. Mervin E. Muller will speak on the purposes and aspirations of IASC; Ivor Francis, on the portability, evaluation, and certification of programs by statistical associations and societies; Tore Dalenius, on international implications of computer privacy; Graham Wilkinson, on language requirements and designs to aid data analysis and computing; Julio Ortuzar, on editing census and survey data using small computers in developing countries, including a review of experience with CONCOR; and, Geoffrey Thomas, on assisting data processing activities in developing countries. [All appear to be quite relevant topics for IASSIST members. Let's hope that papers or proceedings will be available.]

Further information and details can be obtained by writing: International Statistical Institute, 428 Prinses Betrixlann, Voorburg, Netherlands. Annual individual member's dues structure is (1) members from developed countries, $15 per year; and (2) members from developing countries who request a reduced rate, $7 per year.

PUBLICATIONS

The US National Technical Information Service (NTIS) announces the publication of a Directory of Computer Software Applications-Energy. It is 80 page reference source which presents over 300 abstracts of available energy-related reports and computer programs developed from federally sponsored research. The abstracts detail what software is available in fossil, solar, nuclear, geothermal, ocean thermal and other energy areas. The programs range in subject matter from comparative costs to social impact and are offered in several computer languages. You receive complete software program instructions (or a program source) and/or a report for a cost of a paper copy (average cost is $6). The NTIS brochure suggests that the Directory is an ideal source for reviewing available energy software programs for guidance before undertaking expensive program development of your own. The Directory costs $20 and is available in paper copy or microfiche. NTIS Order No. is PB-264 200. Write to NTIS, US Department of Commerce, Springfield, VA 22161.

The analysis of textual data has recently benefitted from the methods and tools which are the result of computer analysis. Editions du CNRS announces a new text on the analysis of textual data: Analyse et validation dans l'etude des données textuelles. The brief description says that "the objective of this work is to put forward a certain number of theoretical perspectives and concrete experiences that present the analysis of texts in the light of data-handling techniques and be a reference to the most valued acquisitions of contemporary linguistics". The book groups 14 papers from French and foreign specialists working in diverse fields of study but all of whom are concerned with the same fundamental questions. The more strictly methodological or theoretical parts presented in each of these contributions are illustrated by reports on particular experiences related to specific branches of knowledge and objectives (archaelogy and history, folklore, literature, anthropology). For further information, write Editions du CNRS, 15 quai Anatole France, 75700 Paris, France; in Canada: Presses de l'université de Montréal, case postale 6128, Montreal 101; USA: SMPF, 14 East 60th Street, New York, NY 10022.

The Social Science Data Archive of the Social Science Library at Yale University announces the revision and updating of its Directory of Data Holdings and Services. Cost is $5.00 and prepayment is requested. Checks should be made payable to Yale University, but sent to the Social Science Data Archive, Social Science Library, Box 1958, Yale Station, New Haven, Connecticut 06520.

Call for Papers

The American Sociological Association (ASA) has issued a call for papers for its next annual meeting (August 1978). Several sessions appear to provide an opportunity for IASSIST members to address matters of interest to both sociologists and data archivists and librarians, in particular, USES OF SECONDARY DATA. Papers are welcome. The deadline is January 10, 1978 for submission of complete papers. The Association's rule is that all participants on the program must be members of ASA. Charles R. Wright, responsible for the session, notes that authors should "make explicit the methodological aspects of their contributions, although an essentially substantive paper that also served as a demonstration of the uses of secondary data would not be inappropriate." Potential contributors should contact Professor Charles R. Wright, Annenberg School of Communications, C05, University of Pennsylvania, Philadelphia, PA 19104.

UPCOMING MEETINGS

IASSIST NORTH AMERICA
february 8–11, 1978

Chicago Meeting -- February, 1978

AGENDA (Revised)

THEME: State of the Art: Perspectives

DATE: February 8-11, 1978

PLACE: Nordic Hills
Itasca, Illinois (limousine service is available from Chicago's O'Hare airport)

PROGRAM COMMITTEE: Tony Falsetto (Chair), Public Archives of Canada
Sheldon Laube, C. M. Leinwond Associates
Richard Roistacher, University of Illinois
Carolyn Geda, ICPSR
Patrick Bova, NORC

PANELS:

1. Documentation
2. Acquisition and Preservation
3. Potentials and Problems of Networking
4. Privacy versus Freedom of Information
5. Alternative Structures for Data Access
6. Software Analysis of Non-Rectangular Files

INDIVIDUALS INTERESTED IN PRESENTING A PAPER AT ANY OF THE SESSIONS SHOULD SEND THE PROPOSED TITLE WITH A 100 WORD ABSTRACT to Carolyn Geda, Chairperson, IASSIST, ICPSR, P. O. Box 1248, Ann Arbor, MI 48106

Workshops and Action Groups will also be scheduled. The schedule for the panels, workshops and action group meetings will be sent to IASSIST members with conference registration forms during November. PLEASE NOTE: individuals who were following the original program theme and who anticipated doing a formal presentation within the Action Group should submit their paper to the Chairperson.

IASSIST / ISA
august 14 – 21, 1978

International Sociological Association
Uppsala-August 14-21, 1978

THEME: International Interchange of Data

PROGRAM COMMITTEE: John DeVries (Chair), Carleton University, Canada
Elliott Avedon, University of Waterloo, Canada
Carolyn Geda, Inter-university Consortium for Political and Social Research, University of Michigan, U.S.A.
Don Harrison, National Archives, Washington, D.C., U.S.A.
Guido Martinotti, University of Turin, Italy
Per Nielsen, Danish Data Archives, Denmark

PANELS:

1. Privacy versus Freedom of Information
   Chair: Guido Martinotti, Archivio dati e programmi per le scienze sociali, via G. Cantoni 4, 20144 Milan, Italy
2. Issues in Comparative Data and Research
   Chair: Elliott Avedon, Department of Recreation, University of Waterloo, Waterloo, Canada
3. Research Problems Associated with Complex Data Bases
   Chair: John DeVries, Department of Sociology, Carleton University, Ottawa, Ontario, Canada K1S 5B6

Papers are solicited. Individuals interested in presenting a paper should contact the chairperson as soon as possible.

Information on the availability of charter flights will be sent to IASSIST members in the near future.

Workshops will be scheduled informally. A Business Meeting will also be scheduled.

POSITION OPENING
U.S. BUREAU OF THE CENSUS
WASHINGTON, D.C.

Archiving and Documentation Specialist (GS 9-11)

The incumbent of this position will serve as an archiving and technical documentation specialist in the Customer Services Branch of the Data User Services Division. The incumbent is responsible for preparing and maintaining an archiving and technical documentation standards manual for publicly distributable computer software, machine-readable data, and microform products. In accordance with the standards manual, the incumbent prepares technical documentation, and related user materials, for products available from the Bureau. Assists and works with other Bureau divisions in developing format specifications most appropriate to meet user needs. Responsible for a broad based archiving plan including operating procedures for copying and preparing extracts and printouts, maintenance of backup copies, maintenance of historical processing steps as appropriate, and a current register of all such products available from the Bureau including description and availability.

Applicants should possess a college degree in business, economics, or sociology and 2-3 years experience in the use of machine-readable data products. Applicants must also possess training, experience, and aptitude for developing documentation, standards and guidelines for preparing and processing machine-readable products. Expertise and proven ability in preparing extensive technical documentation for computer files essential.

For further information, contact:

Larry Carbaugh, Chief
Customer Services Branch
Data User Services Division
U.S. Bureau of the Census
Washington, D.C. 20233
(301) 763-2400

Positions are filled through the Civil Service Commission System.

Equal Opportunity Employer.

INDEX

African Studies Handbook 1:2:32
American National Standards for Bibliographic References 1:2:11
American Sociological Association, Annual Meeting, August 1978 1:4:44
Analyse et validation dans l'étude des données textuelles 1:4:43
Anglo-American Cataloguing Rules II 1:2:11
1:3:9

Ann Arbor Summer Workshops 1:1:18-19
1:2:29

"An Overview of Display Terminals" 1:4:19-20
"A Source Documentation Style Manual" 1:4:32-33
Australian Consortium for Social and Political Research, Inc. 1:2:5,26
Avedon, Elliot M. 1:2:4,9
1:4:7,45

Belgian Archives for the Social Sciences (BASS) 1:3:31
1:4:9

Bellagio Conference on Privacy, Confidentiality, and the Use of Government Micro-Data 1:4:21-23
Bibliographic References, Standards 1:2:11
1:3:9-10
1:4:17

Bonmariage, Joseph 1:1:2,3,9
1:2:3,7
1:3:3,6
1:4:3-4

Book Notices *see* Heim, Kathleen M.
Borman, Lorraine 1:4:38
Bova, Patrick 1:4:44
Boys Town Center for the Study of Youth Development 1:2:31-32
Brewster, Kenneth 1:2:5
Brouwer, Evert 1:1:2
1:2:3
1:3:3
1:4:3

Canadian Secretariat Report 1:1:7
1:2:4-5
1:3:19

Cataloguing *see also* Indexing and Classification 1:3:23-24
1:4:11-18,38

Chairperson's Report 1:1:3
1:2:4-5
1:3:4
1:4:5-6

Chapple, Sharon *see* Henry, Sharon Chapple
Chicago Meeting, February 1978 *see* North American Conference, February 8-11, 1978
Classification Action Group 1:1:12-13
1:2:8-11
1:3:8-10
1:4:9

Cocoa Beach Meeting *see* North American Conference, February 16-20, 1977
Codebook *see* Source Documentation
Committee of European Social Science Data Archives 1:2:25
Confidentiality 1:2:24-25
1:4:20-28
Constitution 1:1:5-7
Crewe, Ivor 1:1:2,4
1:2:3
1:3:3
1:4:3
Danish Data Archive (DDA) 1:4:9
Data Abstract *see* Study Abstract
Data Acquisition Action Group 1:1:10
1:2:7-8
1:3:22
Data Acquisition Policies, Questionnaire 1:4:8
Data Archive Development Action Group 1:1:14
1:2:11-13
1:3:11
Data Archive Registry Action Group 1:1:9
1:2:7
1:3:6-7,20
Data Documentation Action Group 1:1:11
1:2:8
Data Needs Announcements 1:1:20
Data Organization and Management Action Group 1:1:16-18
1:2:15-17
1:3:15-16
Data Organization Registry Form 1:4:5,8,10,34-36
Data Sources 1:4:7
Des Rosiers, Rachel 1:3:19
DeVries, John 1:2:4
1:3:21
1:4:4,45
Directory of Computer Software Applications--Energy 1:4:43
Directory of Data Bases in the Social and Behavioral Sciences 1:3:6-7
1:4:8
Directory of Data Holdings and Sources 1:4:43
Directory of Directories *see* Process-Produced Data Action Group
Directory of Social Sciences Data Centres 1:3:7
Display Terminals 1:4:19-20
Documentation *see also* Data Documentation Action Group, Codebook 1:4:7-8
Dodd, Sue 1:1:12
1:2:8
1:3:7,8-10,23-24
1:4:4,9,11-18,38
DORF *see* Data Organization Registry Form
Drexel Library Quarterly 1:4:37-38
Dues 1:1:3
Edinburgh, Scotland, Meetings, August 16-20, 1976 1:1:3,8
-Agenda 1:1:17-18
see also International Political Science Association
Editor's Note 1:1:3

Employment Opportunities 1:1:20
1:3:38
1:4:46
Essex Summer School in Social Science Data Analysis and Collection, 10th, Colchester, England, July 15-August 26, 1977 1:2:31
European Regional Cartography Meeting, Bergen, September 1-3, 1977 1:3:34-35
Falsetto, Tony 1:3:4,36
1:4:44
Ferguson, Douglas 1:4:37
Flaherty, David 1:4:21-23
Framework for Planning U.S. Federal Statistics 1978-1989 1:4:40
French Version of IASSIST *Newsletter* 1:1:20
Gammell, William 1:1:16
1:4:4
Geda, Carolyn 1:1:2,3,4
1:2:3,4
1:3:3,4,19
1:4:3,5,7,38,44-45
Governance 1:1:6
Grandon, Gary M. 1:4:30-32
Graphic Display of Spatial Data 1:3:32
Grasso, John 1:1:11
1:2:8
Guide to Providing Social Science Data Services 1:2:4,12
1:3:11-12, 16
1:4:8
Handbook of American Resources for African Studies 1:2:32
Hanis, Ed 1:2:3,4,12
1:3:3
1:4:3,6
Harrison, Donald 1:1:10
1:2:4
1:4:4,28-30,45
Hay, Richard Jr. 1:4:38
Heim, Kathleen M. 1:2:21
1:3:25-31
1:4:37-40
Henrichsen, Bjørn 1:1:2-3,18
1:2:3
1:3:3
1:4:3
Henry, Sharon Chapple 1:1:2,4,7
1:2:3-4,9
1:3:3,7
1:4:3,7
INSSIST, International Social Science Information Services and Technologies 1:4:40
Indexing and Classification *see also* Classification Action Group, Subject Headings 1:4:7,8,10
International Association for Statistical Computing 1:4:42
International Committee for Social Sciences Information and Documentation (ICSSD) 1:2:7
International Federation of Data Organizations (IFDO) 1:3:32
1:4:10

International Political Science Association World Congress, Edinburgh, Scotland, August 1976	1:1:3
International Sociological Association World Congress, Uppsala, Sweden, August 1978	1:1:4 1:2:4,33 1:3:38 1:4:5,45
Inter-university Consortium for Political and Social Research (ICPSR)	1:4:9,26-27
-Preservation on Confidentiality	1:4:26-27
-Workshop on Management, Library Control and Use of Computer Readable Information	1:2:29
Kolp, John	1:2:4,6 1:4:4
Lacasse, Pierre	1:1:10 1:2:7 1:3:7,22 1:4:3
Lanphier, C. Michael	1:1:18
Lasko, Lisa	1:1:9 1:2:7 1:3:6-7,20 1:4:4
Laube, Sheldon	1:2:4 1:4:4,44
Laurent, Phillippe	1:2:7
LEAA Research Support Center Machine-Readable Source Documentation System: User's Guide	1:4:33
Leavitt, Michael	1:1:15
Leisure Studies Data Bank (LSDB)	1:4:9
Lucci, York	1:3:25-28
Madron, Thomas	1:1:2-3 1:2:3 1:3:3 1:4:3,5
Martinotti, Guido	1:1:2-3,18 1:2:3 1:3:3 1:4:3,45
Meeting *see* North American, West European, or Sponsor's Name *see also* Upcoming Meetings	
Meetings, Future *see* Upcoming Meetings	
Melton, Michael	1:4:19-20
Membership	1:1:6,7,8 1:2:6 1:3:5 1:4:6
Meyerhoff, Eric	1:3:25-28
Meyriat, Jean	1:2:7 1:3:7
Middendorp, Cees	1:1:2-3,11 1:2:3,8 1:3:3 1:4:3-4,7

Mochmann, Ekkehard 1:1:2-3,12
1:2:3,8
1:3:3,8
1:4:3-4,7

Mora y Araujo, Manuel 1:1:2
1:2:3
1:3:3
1:4:3

Morrison, Greg 1:1:16
1:3:24
1:4:4

Müller, Paul 1:1:15
1:4:4

Nasatir, David 1:1:2,9
1:2:3,7,22-24
1:3:3
1:4:3,38

Nielsen, Per 1:1:2,3,8
1:2:3,4,5
1:3:3
1:4:3,5,7,45

Nijhawan, Naresh 1:1:2,4
1:2:3
1:3:3
1:4:3

Noble, Barbara 1:4:32-33,39-40

Nominating Committee 1:4:4

North American Conference, Cocoa Beach, Florida, February 16-20, 1977 1:1:4
1:2:4,6

North American Conference, Toronto, Canada, May 8-12, 1977 1:2:33
1:3:19-24,passion

North American Conference, Nordic Hills, Itasca, Illinois, February 8-11, 1978 1:2:33
1:3:36-37
1:4:5,44-45

Objectives, IASSIST 1:1:5

"On Work in Progress: Directory of Directories" 1:4:28-30

Open Conference, 1st, Toronto, May 11-12, 1977 1:2:33

Øven, Ørjar 1:2:24-25

Ostrowski, Krzysztof 1:1:2,4
1:2:3
1:3:3
1:4:3,7

Perspectives in Social Science Data Archiving: An Annotated Bibliography 1:2:13

Polish Academy of Sciences, Section on Data Processing and Data Archiving 1:2:27

Privacy *see* Confidentiality

Proceedings, North American Conference, Toronto, Canada, May 8-12, 1977, Abstracts of Papers Presented 1:3:19-24

Process-Produced Data Action Group 1:1:15-16
1:2:13-15,17-21
1:3:13-14
1:4:28-30

Publication Policy 1:3:18
QUANTUM-Social Science History Association Conference, Cologne, August 10-12, 1977 1:3:33
QUANTUM Survey of Research Projects 1:3:31
Radcliffe College Data Resource and Research Center 1:2:27-28
Regional Secretariats 1:1:4
Reviews 1:1:19
Robbin, Alice 1:1:2,3,4,14
  1:2:3,4,11
  1:3:3,11
  1:4:3,4,5,20-28,
Roistacher, Richard 1:2:4
  1:4:32-33,39-40,
Rokkan, Stein 1:2:7
  1:3:25-28
Roper Public Opinion Research Center Reorganization 1:2:28
Rowe, Judith 1:1:2,3,9
  1:2:3,4,12
  1:3:3,5,20
  1:4:3,38
Ruus, Laine 1:1:14
  1:2:11-12
  1:3:11,22
  1:4:3
Salley, Dave L. 1:1:11
  1:2:8
  1:3:23
  1:4:4
Scheuch, Erwin K. 1:1:18
Sharma, Mohan 1:2:8-9
  1:3:8
  1:4:4
S S Data 1:2:4
"State of the Art: Perspectives", Chicago, February, 1978 1:2:33
  1:3:36-37
  1:4:44-45
"Social Science Data in Canada", Toronto, May 8-10, 1977 1:2:33
Source Documentation 1:4:8,16-18,32-3
  -Style Manual 1:4:32-33
  -Title Page 1:4:16-18
Statistical Package for the Social Sciences (SPSS) 1:4:30-32,33
Steering Committee 1:1:2
  1:2:3
  1:3:3
  1:4:3
Steinmetzarchiv 1:4:9
Study Abstract 1:4:9
Study Description 1:4:9
Subject Headings for Data Files 1:3:7
Summer Training Seminar in Cross-National Data Analysis, Vienna, July 3-22, 1977 1:2:29-30
Tannenbaum, Eric 1:1:16-17
  1:4:4

Taylor, Marcia 1:1:3,10
1:2:7
1:3:17
1:4:3
"The Issue of Confidential Data: The Need for Formulation of Policy by the Data Archive and Library" 1:4:20-28
"Titles: The Emerging Priority in Bringing Bibliographic Control to Social Science Machine-Readable Data Files [MRDF]" 1:4:11-18
Toronto Meeting *see* North American Conference, May 8-12, 1977
Treasurer's Report 1:4:6
UNESCO Directory of Data Services 1:2:7
1:3:6
UNISIST 1:4:41
University of Wisconsin Data and Program Library Service Authorization Form 1:4:27-28
United States Secretariat Report 1:1:9
1:2:6
1:3:5,20
Upcoming Meetings 1:1:20
1:2:4,33
1:3:33-37
1:4:44-45
Uppsala Meeting, August, 1978 *see* International Sociological Association
"Using SPSS MULT RESPONSE to Generate Filtered Marginals" 1:4:30-32
West European IASSIST Action Group Workshop, Copenhagen, June 26-29, 1977 1:3:33
1:4:5,7-10
West European Secretariat Report 1:1:8
1:2:4-5
White, Howard D. 1:3:29-31
1:4:37
Women-Data Archives 1:2:27-28
Working Manual for Cataloguing Machine-Readable Data Files 1:2:9
Workshop in Data Management, Data Library Activities, and Data User Services, Ann Arbor, Michigan, July 7-16, 1976 1:1:18-19
Workshop on Management, Library Control and Use of Computer Readable Information, Ann Arbor, Michigan, July 25-August 5, 1977 1:2:29
Yale University. Social Science Data Archive 1:4:43

iassist

MEMBERSHIP APPLICATION

NAME ______________________________

MAILING ADDRESS ______________________________

____

INSTITUTIONAL AFFILIATION ______________________________

____

TELEPHONE NUMBER ______________

Date ______________ Amount Enclosed ______________

Membership fees for calendar year 1977

Individual: Regular $15 ____ Student $5 ____

Institutional (two individual memberships): $35 ____

Charter Individual Membership (three years): $100 ____

Institutional Subscription: $25 ____

Payment Enclosed ________ (amount)

Make check or money order payable to IASSIST. Send payment to:

Judith S. Rowe
PRINCETON UNIVERSITY COMPUTER CENTER
87 Prospect Avenue
Princeton, N. J. 08540

For individual memberships, indicate interest in participation in one of the following action groups:

DATA ARCHIVE REGISTRY ____
DATA ARCHIVE DEVELOPMENT ____
DATA ACQUISITION ____

DATA ORGANIZATION AND MANAGEMENT ____
CLASSIFICATION ____
DOCUMENTATION ____
PROCESS-PRODUCED DATA ____

Membership in the IASSIST includes a subscription to the IASSIST *Newsletter* and a subscription to *S S Data, A Newsletter of Social Science Archival Acquisitions*, each to appear four times a year. In addition to these subscriptions, the IASSIST membership affords the opportunity to participate at both the international and regional levels.

- -

(for individual's file)

I applied for membership in IASSIST. Dues paid (amount) ____ Date ____________

iassist

PRINCETON UNIVERSITY
COMPUTER CENTER
87 Prospect Avenue
Princeton, NJ 08540

Non-Profit Org.
U. S. Postage
PAID
Permit No. 179
Princeton, N.J. 0854